

Copyright

by

David Lee Cramer Jr.

2013

**The Dissertation Committee for David Lee Cramer Jr. Certifies that this is the  
approved version of the following dissertation:**

**Design, Synthesis, and Calorimetric Studies on Protein-Ligand Interactions:  
Apolar Surface Area, Conformational Constraints, and Application of the Topliss  
Decision Tree**

**Committee:**

---

Stephen F. Martin, Supervisor

---

Jennifer Brodbelt

---

Dionicio Siegel

---

Eric V. Anslyn

---

Sean Kerwin

---

**Design, Synthesis, and Calorimetric Studies on Protein-Ligand  
Interactions: Apolar Surface Area, Conformational Constraints, and  
Application of the Topliss Decision Tree**

**by**

**David Lee Cramer Jr., B.S.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**April 2013**

## **Dedication**

I would like to dedicate this work to my wife Natania, and to our son Morrissey.



## Acknowledgements

I would like to thank Professor Stephen F. Martin for all of his guidance and mentorship throughout my graduate research studies. I am grateful for his support of my research, and for his committal to developing my critical thinking skills that are essential to becoming a diligent scientist. I must also acknowledge my undergraduate research advisor Professor Jonathan S. Lindsey and my mentor Dr. Masa Taniguchi for encouraging me to pursue my interests in chemistry. I would like to give a special thanks to Dr. James Myslinski, who proofed and edited this dissertation. James was also my steady lab mate and fellow colleague throughout all of the highs and lows of my graduate research. I would like to thank Dr. John Clements for performing all of the structural analysis for the ligands that I worked on, as well for editing parts of this text. For their friendship and encouragement, I want to acknowledge Brett Granger, Alex Goodnough and Caleb Hethcox for their assistance throughout my final year. I also want to thank Mike O’Keffe and John DeLorbe for their help and company when I was just starting out my graduate studies.

I also owe a great amount of gratitude to my wife, Natania Morrissey, who I met in Austin during my second year of graduate research. She has been the most supportive figure of my life, and for that I will be forever indebted to her.

**Design, Synthesis, and Calorimetric Studies on Protein-Ligand  
Interactions: Apolar Surface Area, Conformational Constraints, and  
Application of the Topliss Decision Tree**

David Lee Cramer Jr., Ph. D.

The University of Texas at Austin, 2013

Supervisor: Stephen F. Martin

A preorganised amino acid derivative containing a cyclopropyl constraint was designed to orient an amino acid into its bound conformation. This constrained mimic was determined by ITC to be equally potent to the native Phe derivative. It was found that a more favorable enthalpy of binding was compensated by an equally unfavorable entropy compared to the native ligand. In order to properly ascertain the effects of the cyclopropane constraint, a flexible control containing the same number of heavy atoms was synthesized and tested, and it was found to be at least 200 fold *less* potent than the constrained analog. However, without structural data of the flexible control, it is difficult to infer if the differences in ligand binding affinity arose from the ligand constraint or some other unknown complexity to binding.

We studied the thermodynamic and structural effects of modifying alkyl chains of *n*-alka(e)nol and phenylalka(e)nol binders to MUP-I by both the removal of a rotor via deletion of a methylene unit and restriction of a rotor via the installation of an internal olefin. In general, we observed that a similar thermodynamic signature accompanies modifications for both the *n*-alka(e)nol and phenylalka(e)nol ligands: A favorable

$T\Delta\Delta S^{\circ}_{\text{obs}}$  is compensated by an unfavorable  $T\Delta\Delta H^{\circ}_{\text{obs}}$  such that  $T\Delta\Delta G^{\circ}_{\text{obs}}$  for both removal of a methylene and insertion of an internal olefin are unfavorable and equipotent, respectively. The insertion of an internal olefin into an alkyl chain led to significantly more favorable entropies than does methylene removal, yet enthalpy-entropy compensation leads to nearly equipotent binding energetics. However, we did find a strong correlation between  $\Delta H^{\circ}_{\text{obs}}$  and buried apolar Connolly Surface Area (CSA).

The intrinsic free energies of introducing an internal olefin into the *n*-alkanols and phenylalkanols differ markedly from the observed data. It was observed that intrinsic affinities are more favorable than the observable because a favorable  $T\Delta\Delta S^{\circ}_{\text{int}}$  *dominates* an unfavorable  $\Delta\Delta H^{\circ}_{\text{int}}$ . Also, we discovered that the intrinsic entropies of inserting an internal olefin are nearly double that of removing a methylene group, suggesting that the insertion of an internal olefin results in the restriction of *more* C-C rotors.

We have shown through ITC analysis that the added substituents probed in this study provided binding increases to our Grb2 SH2 ligands as expected, but that the thermodynamic driving force of binding affinities depended greatly upon the specific nature and flexible mobility of the ligands in the binding pocket. Through a combination of X-ray and ITC studies it was shown that ligands containing rigid and aromatic functional groups bound with a higher  $\Delta H^{\circ}$  than the more flexible alkyl ligands, and that this effect correlates well with more direct vdW contacts made in the pocket.

Finally, we described a case study where a strict adherence to the Topliss operational schemes led to an expedient development of novel MUP-I binding analogs. The validity of the schemes was also depicted through the synthesis and testing of ligands that were correctly predicted to be weaker/equipotent to the starting ligand. Of important

note is that the degree to which the schemes led to affinity boost depended greatly on the starting potency of the initial compound.

## Table of Contents

<b>Chapter 1 Strategies of Drug Design: Are they Predictable, Explainable? .....</b>	<b>1</b>
1.1.1 Strategies of Drug Design: Are they Predictable, Explainable? ....	1
1.2 Introduction.....	1
1.3 Binding Energetics of Protein Ligand Interactions. ....	3
1.4 Intrinsic analyses.....	5
1.5 Monitoring Protein Ligand Interactions. ....	6
1.6 Ligand Preorganization and Conformational Constraints.....	12
1.7 Model Systems for Studying Ligand Preorganization.....	17
1.8 Olefins as a rigid constraint .....	27
1.9 Cyclopropyl Constraints .....	31
1.10 Studies of Ligand Binding Affinity Increases Driven by Added Hydrophobicity. ....	41
1.11 Studies on the Thermodynamic Nature of Hydrophobic Effects.	49
1.12 Topliss Operational Schemes. ....	64
1.13 Conclusions. ....	67
<b>Chapter 2 Thermodynamic and Structural Evaluation Preorganizational Constraints for MUP-1 Ligands .....</b>	<b>70</b>
2.1 Mouse Major Urinary Protein I (MUP-I).....	71

<b>2.2</b>	<b>Ligand Design .....</b>	<b>73</b>
<b>2.3</b>	<b>Ligand Synthesis.....</b>	<b>75</b>
<b>2.4</b>	<b>Thermodynamic Studies .....</b>	<b>78</b>
<b>2.5</b>	<b>Summary.....</b>	<b>80</b>
<b>2.6</b>	<b>The Binding Thermodynamics of Phenyl Alkanols .....</b>	<b>81</b>
<b>2.7</b>	<b>Developing the ITC protocol .....</b>	<b>84</b>
<b>2.8</b>	<b>The Utility of the Internal Olefin to Restrict Rotors.....</b>	<b>87</b>
<b>2.9</b>	<b>Ligand synthesis.....</b>	<b>92</b>
<b>2.10</b>	<b>Thermodynamic Binding Studies of Alkyl and Phenyl Alka(e)nols. .....</b>	<b>93</b>
<b>2.11</b>	<b>Observed Thermodynamics of Binding: Insertion of Internal Olefin. ....</b>	<b>101</b>
<b>2.12</b>	<b>Crystal Structures.....</b>	<b>113</b>
	Complex.....	121
	Average All-Atom B-factor .....	121
	(Å <sup>2</sup> ) <sup>[a]</sup> .....	121
	No. bound .....	121
	Water <sup>[d]</sup> .....	121
<b>2.13</b>	<b>Summary and Conclusion .....</b>	<b>124</b>
 <b>Chapter 3 Studies of Increasing Ligand Hydrophobicity to a Known Protein</b>		
<b>Binder</b>	<b>.....</b>	<b>128</b>

3.1	Introduction.....	128
3.2	Grb2 Adaptor Protein.....	128
3.3	Experiment Design .....	130
3.4	Investigation of an Appropriate Carbon Linker .....	133
3.5	Isothermal Titration Calorimetry .....	137
3.6	Structural Studies of the pY+3 ligands of Varying Linker Length n... .....	140
3.7	Summary.....	143
3.8	Substituent Effects at the pY+3 Derived Tripeptides. ....	144
3.9	Thermodynamic studies.....	146
3.10	Temperature Dependent Studies. ....	152
3.11	Structural Studies.....	157
3.12	Phenyl Substituent effects.....	160
3.13	Exploring Leveling Effects .....	162
3.14	Conclusions .....	166
 <b>Chapter 4 Topliss Approach to the Binding Affinity Enhancement of MUP-I</b>		
	<b>Binders. ....</b>	<b>169</b>
4.1	Introduction.....	169
4.2	Experimental Design .....	170

4.3	Ligand Synthesis.....	174
4.4	Effect of Flourination of Alkyl Bonds. ....	181
4.5	Structural Characterization .....	183
4.6	A possible dependence of the Topliss schemes on initial analog potency. ....	190
	Conclusion.....	192
Chapter 5 Experimental.....		194
5.1	General .....	194
5.2	Procedures.....	195
References.....		301



## List of Tables

TABLE 1.6.1 BARTLETT'S PDZ-BINDING MACROCYCLES.....	23
TABLE 1.6.2 INVESTIGATION INTO BENZOATE-BRIDGED MACROCYCLIC CONSTRAINTS BY SPALLER.....	25
TABLE 1.6.3 INVESTIGATION INTO GLUTAMATE-BRIDGED MACROCYCLIC CONSTRAINTS BY SPALLER.....	26
<b>TABLE 1.8.1. SRC SH2 BINDING STUDIES.<sup>A</sup></b> .....	34
TABLE 1.8.2. GRB2 SH2 BINDING LIGANDS.....	36
TABLE 1.8.3. GRB2 SH2 EXPANDED LIGANDS.A .....	38
TABLE 1.9.1 PY+1 LIGAND STUDY THAT BOUND TO THE GRB2 SH2 DOMAIN.....	44
TABLE 1.9.2 BINDING STUDIES OF LINEAR PY+1 LIGANDS WITH THE GRB2-SH2 DOMAIN IN HEPES AT 298 K.....	46
TABLE 1.10.1 KLEBE LIGAND WITH MODIFIED ALKYL SURFACE AREA.....	53
TABLE 1.10.2 HOMANS STUDY OF INCREASING HYDROPHOBICITY OF PYRAZINE DERIVATIVES.....	55
<b>TABLE 1.10.3</b> HOMANS LIGAND ANALYSIS WITH MUP-I UTILIZING ALKANOLS.....	56
TABLE 1.10.4 ENBERTS STUDY OF INCREASING METHYLENES TO TRYPSIN BINDERS.....	58
TABLE 1.10.5 SPALLER STUDIES ON SUBSTITUENT EFFECTS.....	63
TABLE 2.4.1. ITC BINDING STUDIES PHE DERIVATIVES WITH MUP-I, AT 298 K IN PBS BUFFER AT PH 7.4 .....	80
TABLE 2.10.1. ITC BINDING STUDIES ALKYL ALKANOLS WITH MUP, AT 298 K IN PBS BUFFER, PH 7.4, PERFORMED BY RICHARD MALHAM.....	95
<b>TABLE 2.10.2.</b> ITC BINDING STUDIES PHENYL ALK(E)NOLS WITH MUP, AT 298 K IN PBS BUFFER, PH 7.4 <sup>A</sup> .....	96
<b>TABLE 2.10.3.</b> DIFFERENCE IN OBSERVED ITC FOR THE REMOVAL OF A METHYLENE GROUP IN ALKYL ALKANOLS.....	98

TABLE 2.10.4. DIFFERENCE IN OBSERVED ITC FOR THE REMOVAL OF A METHYLENE GROUP IN PHENYL ALKANOLS.....	99
TABLE 2.11.1 DIFFERENCE IN OBSERVED ITC FOR THE INSERTION OF A C-C DOUBLE BOND INTO ALKYL ALKANOLS AT 298 K IN PBS BUFFER AT PH 7.4.....	101
<b>TABLE 2.11.2</b> DIFFERENCE IN OBSERVED ITC FOR THE INSERTION OF A C-C DOUBLE BOND INTO ALKYL ALKANOLS.....	102
<b>TABLE 2.11.3</b> INTRINSIC DIFFERENCE IN OBSERVED ITC FOR THE REMOVAL OF A METHYLENE GROUP IN ALKYL ALKANOLS. WORK DONE BY RICHARD MALHAM....	108
<b>TABLE 2.11.4</b> INTRINSIC DIFFERENCE IN OBSERVED ITC FOR THE REMOVAL OF A METHYLENE GROUP IN PHENYL ALKANOLS.....	109
TABLE 2.11.5 INTRINSIC DIFFERENCE IN INTRINSIC ITC FOR THE INSERTION OF A C-C DOUBLE BOND INTO THE ALKYL ALKANOLS, DONE BY RICHARD MALHAM.....	111
TABLE 2.11.6. INTRINSIC DIFFERENCE IN INTRINSIC ITC FOR THE INSERTION OF A C-C DOUBLE BOND INTO THE PHENYLALKANOL.....	112
TABLE 2.12.1. STRUCTURAL FEATURES OF THE COMPLEXES. <sup>A</sup> .....	121
TABLE 2.12.2. CHANGE IN NUMBER OF BOUND WATERS AVERAGED PER MODIFICATION. ....	122
TABLE 3.5.1 ITC BINDING STUDIES WITH GRB2 SH2, AT 298 K IN HEPES BUFFER, PH 7.5. <sup>A</sup> ITC PERFORMED BY DR. BO CHENG. ....	139
TABLE 3.6.1 VAN DER WALLS CONTACTS FOR LIGANDS 3.07 AND 3.08. ....	143
<b>TABLE 3.9.1</b> ITC BINDING STUDIES WITH GRB2 SH2, AT 298 K IN HEPES BUFFER, PH 7.5....	148
TABLE 3.9.2 ITC BINDING STUDIES WITH GRB2 SH2. ....	151
TABLE 3.10.1 ALL ACQUIRED VALUES FOR $\Delta CP^\circ$ .....	156
TABLE 3.11.1 TABULATED VDW CONTACTS.....	160
TABLE 3.12.1 ITC BINDING STUDIES WITH GRB2 SH2, AT 298 K IN PBS BUFFER, PH 7.5. <sup>A</sup> SYNTHESIZED AND TESTED BY DR. J. TIAN.....	162
TABLE 4.2.1 BINDING OF PHENYL ALKANOLS WITH MUP-I AT 298 K IN PBS. <sup>A</sup> .....	171

TABLE 4.3.1 ITC ANALYSIS OF TOPLISS-DERIVED ANALOGS FOR MUP-I AT PH 7.4 IN PBS. <sup>A</sup>	
.....	178
TABLE 4.4.1. ITC OF EXPANDED SERIES IN PBS AT PH 7.4 AT 298 K. <sup>A</sup>	183
TABLE 4.5.1 STRUCTURAL CHARACTERISTIC THAT WERE OBSERVED FROM THE PROTEIN	
BOUND COMPLEXES OF C4PH, C4PME AND C4PCF <sub>3</sub> .....	186

.

## List of Figures

FIGURE 1.1.1 ATTRITION RATES FOR DRUG LEADS SORTED BY YEAR AND CAUSE.....	2
<b>FIGURE 1.4.1.</b> TYPICAL ITC THERMOGRAPH.....	9
<b>FIGURE 1.4.2</b> GRAPHS OF ITC ISOTHERMS SHOWING THE DIFFERENCES IN <i>C</i> VALUES.....	11
FIGURE 1.5.1. PROPOSED BENEFITS OF LIGAND PREORGANIZATION .....	13
FIGURE 1.5.2 JENKS STUDY ON ROTOR RESTRICTION.....	14
FIGURE 1.5.3. MODES OF PEPTIDE PREORGANIZATION. ....	16
FIGURE 1.5.4 PROPER SCISSION AND EXCISION CONTROLS.....	16
FIGURE 1.6.1 BARTLETT'S MACROCYCLIC ANALOGUES.....	18
FIGURE 1.6.2 BARTLETT'S PNECILOPSIN-BINDING LIGANDS.....	20
FIGURE 1.6.3 FREIDINGER STUDY ON LH-RH.....	21
FIGURE 1.7.1. WILLIAMS STUDY OF RISOTCETIN INHIBITORS.....	28
<b>FIGURE 1.7.2.</b> KNOWLES OLEFIN STUDIES.....	29
FIGURE 1.7.3. VERDINE OLEFIN STUDIES.....	30
FIGURE 1.8.1 EARLY MODEL OF CYCLOPROPYL PEPTIDE MIMETICS.....	32
FIGURE 1.8.2. GROUPING OF POPULATION OF PROPOSED SOLUTION STATE STRUCTURES OF BOTH CONSTRAINED AND FLEXIBLE LIGANDS.....	40
<b>FIGURE 1.9.1.</b> AFFINITIES OF PY+1 RESIDUES.....	42
<b>FIGURE 1.9.2.</b> PY+3 DERIVED LIGANDS.....	48
FIGURE 1.10.1. DIEDERICH HYDROPHOBICITY STUDIES.....	50
FIGURE 1.10.2. ROSS LIGAND ANALYSIS.....	51
<b>FIGURE 1.10.3.</b> KLEBE LIGANDS.....	52
<b>FIGURE 1.10.4.</b> WATER BINDING NETWORKS.....	59
FIGURE 1.10.5. FREIRE LIGANDS WHERE BINDING CAVITY IS FILLED VIA METHYL SUBSTITUTION.....	61
FIGURE 1.11.1. TOPLISS OPERATIONAL SCHEME.....	65
FIGURE 1.11.2. BERKMAN LIGANDS.....	67

FIGURE 2.1.1 STRUCTURE OF MUP-I DISPLAYING THE B BARREL BINDING CAVITY .....	71
FIGURE 2.1.2 PROTEIN-LIGAND ASSOCIATION A LIGAND AND A “DEWETTED” PROTEIN BINDING DOMAIN. (A) GENERALLY ACCEPTED HYDRATION OF A PROTEIN AND (B) THE “DEWETTED” HYDRATION STATE. ....	72
FIGURE 2.2.1 MUP-I LIGAND DESIGN FEATURING (A) THE CRYSTAL STRUCTURE OF NHACPHEOME IN THE PROTEIN BOUND COMPLEX AND (B) AND OVERLAY OF THIS STRUCTURE WITH THE PROPOSED CONSTRAINED MIMIC. ....	74
FIGURE 2.2.2 LIGAND DESIGN FOR STUDIES WITH MUP-I. ....	75
<b>FIGURE 2.6.1.</b> ABOVE: BINDING POSE OF NONANOL (A) AND DECANOL (B) WHEREIN IT ADOPTS A TURNED CONFIGURATION (SHOWN AS A DASH). BELOW: CYCLOHEXYL (#) AND PHENYL (#) ALCOHOLS THAT WERE PROPOSED TO MIMIC THE TURN.....	83
FIGURE 2.7.1 EXTINCTION COEFFICIENT CALCULATION .....	86
FIGURE 2.7.2. INDIVIDUAL RUN DATA ANALYSIS OF MUP-I LIGANDS AT 298 K IN PBS BUFFER AT PH 7.4. ....	87
FIGURE 2.8.1 PROPOSED RESTRICTOR STUDY .....	89
FIGURE 2.12.2 COMPLEXES OF MUP-I WITH THE <i>N</i> -ALKA(E)NOLS OBTAINED AT LEEDS UNIVERSITY. OXYGEN AND NITROGEN ATOMS ARE COLORED RED AND BLUE, RESPECTIVELY. <sup>A</sup> .....	115
FIGURE 2.12.3 COMPLEXES OF MUP-I WITH THE <i>N</i> -ALKA(E)NOLS. <sup>A</sup> .....	116
FIGURE 2.12.5 COMPLEXES OF MUP-I WITH THE <i>N</i> -ALKA(E)NOLS. <sup>A</sup> .....	117
FIGURE 3.2.1 X-RAY OF THE GRB2 SH2 DOMAIN BOUND TO A HEPTAPEPTIDE. ....	129
FIGURE 3.2.2 BIOLOGICAL ROLE OF GRB2.....	130
FIGURE 3.3.1 NOVARTIS LIGANDS CONTAINING INCREASE APOLAR SURFACE AREA.....	131
FIGURE 3.3.2 KNOWN MODEL OF PY+3 SUSBTITUENTS INTERACTING WITH AN EXTENDED HYDROPHOBIC PATCH OF THE GRB2 SH2 DOMAIN. ....	132
FIGURE 3.3.3 LIGAND FRAMEWORK FOR STUDIES ON HYDROPHOBICITY. ....	133
FIGURE 3.4.1 TRIPEPTIDE SKELETON FOR THE HYDROPHOBICITY STUDIES.....	134
FIGURE 3.6.1 COCRYSTAL OF THE GRB2 SH2 DOMAIN (SURFACE). <sup>[A]</sup> .....	142

FIGURE 3.8.1 PY+3 SUBSTITUENTS WHERE R IS VARIED ON A THREE CARBON LINKER.....	145
FIGURE 3.10.1 $\Delta H^\circ$ OF BINDING TO THE GRB2 SH2 DOMAIN AS A FUNCTION OF TEMPERATURE.....	152
FIGURE 3.10.2 $\Delta H^\circ$ OF BINDING TO THE GRB2 SH2 DOMAIN AS A FUNCTION OF TEMPERATURE FOR 3.08.....	153
FIGURE 3.10.3 $\Delta H^\circ$ OF BINDING TO THE GRB2 SH2 DOMAIN AS A FUNCTION OF TEMPERATURE FOR 3.28.....	153
FIGURE 3.10.4 $\Delta H^\circ$ OF BINDING TO THE GRB2 SH2 DOMAIN AS A FUNCTION OF TEMPERATURE FOR 3.13.....	154
FIGURE 3.11.1 COCRYSTAL OF THE GRB2 SH2 DOMAIN (SURFACE). <sup>[A]</sup> .....	159
FIGURE 4.2.1 POSE OF 6-PHENYLHEXANOL (4.00) (TABLE 4.1, $N = 4$ ) IN THE BINDING COMPLEX OF MUP-I.....	172
FIGURE 4.2.2 TOPLISS DECISION TREE FROM AN UNSUBSTITUTED PHENYL RING OF A KNOWN DRUG LEAD.....	174
FIGURE 4.3.1. TOPLISS ANALYSIS.....	177
<b>FIGURE 4.5.1</b> STRUCTURAL COMPLEXES OF MUP-I WITH THE TOPLISS DERIVED ANALOGS <b>C4PH, C4PME, AND C4PCF3</b> WERE OBTAINED TO A RESOLUTION OF <2.0 Å RESEARCH BY RESEARCHER DR. JOHN CLEMENTS. OXYGEN AND NITROGEN ATOMS ARE COLORED RED AND BLUE, RESPECTIVELY. CARBON ATOMS BELONGING TO THE PROTEIN (LINE) ARE COLORED GREEN WHILE THOSE BELONGING TO THE LIGANDS (STICKS) ARE COLORED MAGENTA OR CYAN, THE LATTER SHOWN FOR COMPLEXES IN WHICH THE BOUND LIGANDS OCCUPY MORE THAN ONE POSE. PROTEIN-LIGAND DIRECT CONTACTS AND THOSE MEDIATED BY WATER MOLECULES (RED SPHERES) ARE INDICATED WITH DASHED, BLACK LINES. (A) COMPLEX WITH <b>C4PH</b> . (B) COMPLEX WITH <b>C4PME</b> . (C) COMPLEX WITH <b>C4PCF<sub>3</sub></b> .....	185
FIGURE 4.6.1 INCREASE IN THE BINDING AFFINITY OF EACH ANALOG RELATIVE TO THE KNOWN OPTIMAL INCREASE AT THE ONSET OF STUDY. HEREIN CALLED THE RELATIVE PERCENT INCREASE OR RPI. ....	192

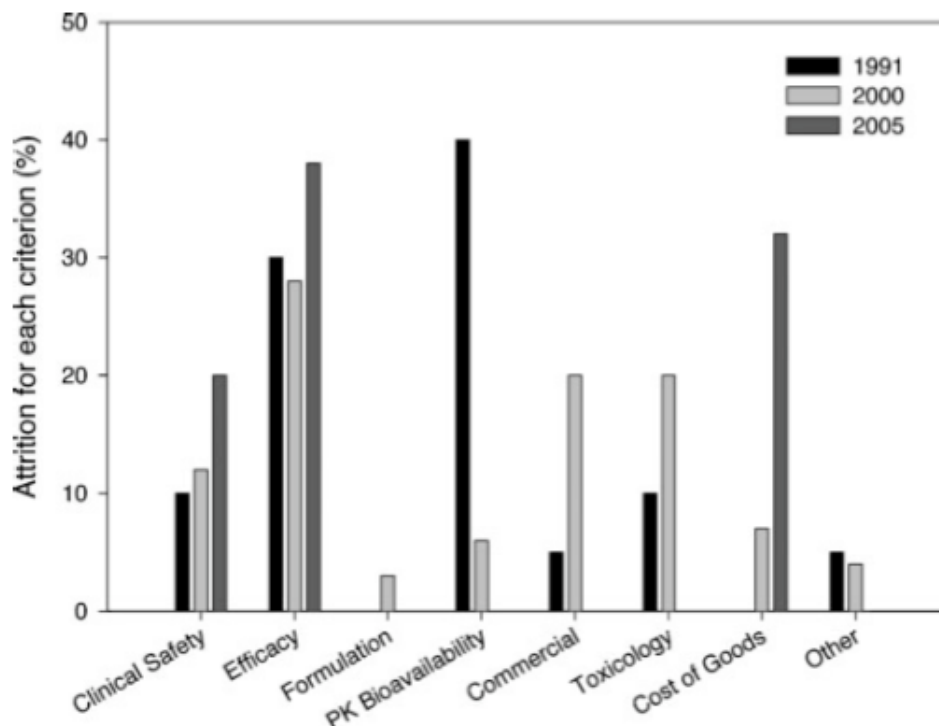
**1.1 Introduction**

Despite the recent advances and improvements to numerous scoring functions and programs that are used to predict a protein's affinity of small molecules, there has yet to be developed a reliable computational method for use in drug discovery.<sup>i,ii</sup> Two of the main roadblocks that stand in the way of achieving this goal are: (1) explicitly accounting for solvent during the binding process and (2) accurately obtaining the energy minima of protein-ligand (solute-solute) interactions necessary to predict the conformation of the ligand in the bound structure.<sup>iii</sup> An alternate route to making early decisions in drug desing is the use of high throughput screening (HTS) of synthetic intermediates. This technology involves the brute force screening of vast biological assays to identify early drug leads. HTS originated as a technique for pharmaceutical companies to screen compounds internally. After some tepid growth in the 1980s, its use erupted in the 1990s so that by the beginning of the new millennium, weekly throughput capability of HTS increased six-fold to 150,000 compounds per week per target.<sup>iii,iv</sup>

One of the fundamental reasons that HTS caught on so strongly in the pharmaceutical industry is that it fostered easy access to screening of vast libraries of compounds for hits rather than relying on human intuition to select candidates, thus ensuring that no blockbuster drug was left on a dusty shelf.<sup>iv,v</sup> Fortuitously, additional benefits of this process arose when it was observed that the early screening of drug leads correlated with a decline in the attrition of compounds due to bioavailability and

pharmacokinetics ( Figure 1.1.1).<sup>v</sup> Along with HTS came a technology boom in the area of practical and user-friendly screening techniques such as time-resolved fluorescence spectroscopy techniques that further enhanced its performance making it more useful.<sup>vi,vii</sup>

**Figure 1.1.1** Attrition rates for drug leads sorted by year and cause.<sup>viii</sup>



However, as the resources and capital dedicated to HTS have increased, genuine discoveries, as measured by investigatory new drug submission (IND), remained relatively unchanged.<sup>ix</sup> It has thus been suggested that drug discovery as an industry should go back to the basics of unraveling the science to develop early leads as opposed to adhering to brute force methods.<sup>ix</sup> On this front, many of the fundamental design principles and strategies used in drug discovery have come into question.<sup>x,xi,xii</sup> Thus, a better understanding of bimolecular interactions would be beneficial to both drug discovery and to the scoring functions aimed at predicting structure-based binding affinity. Toward this goal, it is of immediate importance to ask if these fundamental



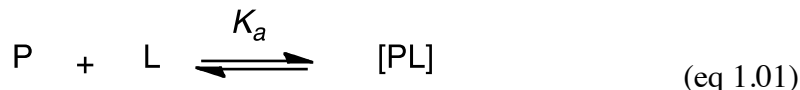
strategies in drug design can be explained rationally by correlating structure with thermodynamic parameters. Second, one must also ask if the structural and energetic effects of these structural modifications can be accurately predicted. Therefore, we pursued a centralized goal of improving the understanding of protein-ligand interactions with the hope and goal of improving how they can be implemented in the design of therapeutics and predicted *in silico*.

This chapter is dedicated to covering in detail some of the known strategies that are thought to increase binding affinity and that are used at the early stages of drug discovery. Of key significance will be a discussion on if these strategies are well explained and predictable. We will focus on significant contributions to this field that decompose protein-ligand interactions to provide the entropy and enthalpy of binding and to explore for correlating thermodynamics and structural data. We will also examine the literature for studies that present a well-designed and systematic experiment that includes the use of proper controls. Ultimately, we hope to conclude this chapter with an outline of what is lacking in the field and what needs to be accomplished to make fundamental discoveries that shall benefit drug development in the pharmaceutical industry.

## **1.2 Binding Energetics of Protein Ligand Interactions**

Given paucity of knowledge on the basic fundamental properties by which protein-ligand interactions are governed, it is important to develop a comprehensive approach to understand binding affinity [ $K_a$ ] in biomolecular interactions.<sup>xiii</sup> Since the ligand binding affinity [ $K_a$ ] is related to the potency of pharmaceutical compounds, we believe that it is first necessary to improve upon the understanding of these interactions

before one can accurately predict the protein-ligand binding affinities. As shown by eq 1.01,  $K_a$  is the equilibrium constant between protein (P) and ligand (L) in solution and the nascent protein-ligand complex (PL): The larger the value of  $K_a$ , then the stronger the binding. This value can also be described as  $K_d$ , which is simply the reciprocal of  $K_a$ .



$$K_a = \frac{[PL]}{[P][L]} = 1/K_d \quad (\text{eq 1.02})$$

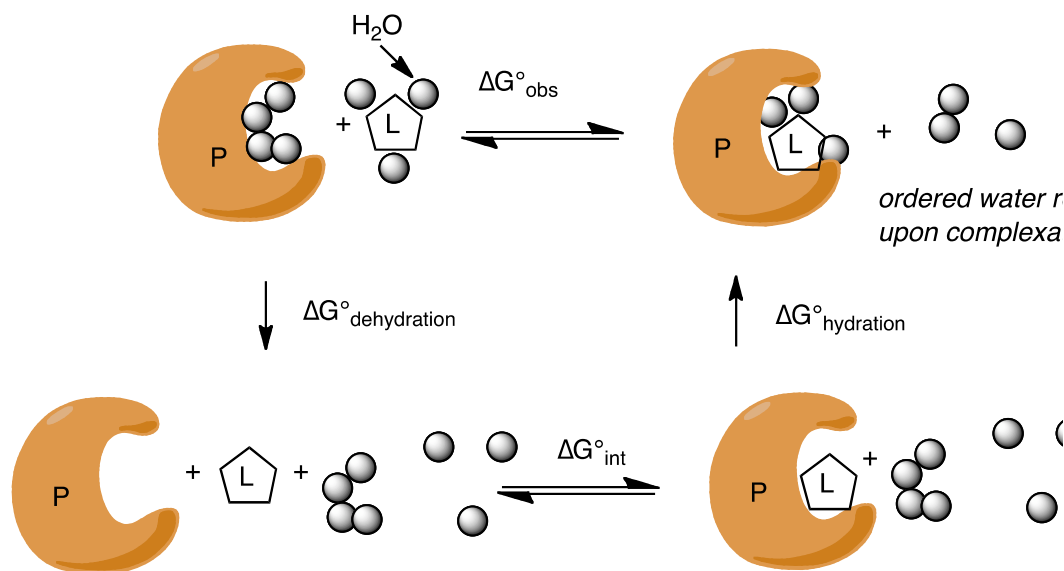
The binding affinity ( $K_a$ ) is proportional to the Gibbs free energy of the binding process,  $\Delta G^\circ$ , given by eq 1.03, which can be further decomposed into its thermodynamic components  $\Delta H^\circ$  and  $\Delta S^\circ$  (eq. 1.01), which are both of critical importance with studies attempting to correlate ligand modifications with individual thermodynamic parameters.

$$\Delta G_{\text{obs}}^\circ = -RT \ln(K_a) \quad (\text{eq 1.03})$$

$$\Delta G_{\text{obs}}^\circ = \Delta H_{\text{obs}}^\circ - T \Delta S_{\text{obs}}^\circ \quad (\text{eq 1.04})$$

Although a complete understanding of biomolecular associations remains a holy grail in the field of protein-ligand interactions, *generally* such processes are understood depicted by the Born-Haber cycle (Scheme 1.2.1).<sup>xiv</sup> Accordingly, the Born-Haber scheme dictates that for binding to occur, both the unbound protein and the ligand in solution must first dehydrate. Attractive forces between the protein and ligand lead to the solute-solute complex, which is called the *intrinsic* binding parameter,  $\Delta G_{\text{int}}^\circ$ . The resulting protein-ligand complex is then *resolvated* to give the *observed* complex and solvent molecules that remain after rehydration are expelled into bulk solvent.

**Scheme 1.2.1.** Born Haber Scheme.



### 1.3 Intrinsic Analyses

All energetic consequences of water organization shown in the Born-Haber cycle are accounted for in the *observable* energetics that are detected by instrumentation. Thus, calculating the *intrinsic* parameters is a nearly impossible task unless one can accurately account for the hydration parameters of the unbound protein, the free ligand, and the bound protein-ligand complex (eq 1.05).

$$\Delta G^\circ_{\text{obs}} = \Delta G^\circ_{\text{int}} - \Delta G^\circ_{\text{hyd,L+P}} + \Delta G^\circ_{\text{hyd,C}} \quad (\text{eq 1.05})$$

To address this problem, it has been proposed that for a set of two *closely* related compounds 1 and 2, one may make the simplifying assumption that differences in  $\Delta G^\circ_{\text{hyd,L+P}}$  and  $\Delta G^\circ_{\text{hyd,C}}$  are negligible, so they can be removed from the equation when calculating the differences in intrinsic parameters (shown in eq 1.06).<sup>xv</sup> However, the main caveat with this assumption is that water molecules in the bound complex are often difficult to discern, and therefore, the hydration parameters of the bound complex may

not be known with certainty. This technique will be discussed in further detail in Chapter 2 of this thesis.

$$\Delta\Delta G^{\circ}_{\text{int},2-1} = \Delta G^{\circ}_{\text{int},2} - \Delta G^{\circ}_{\text{int},1} = \quad (\text{eq. 1.06})$$

$$(\Delta G^{\circ}_{\text{obs},2} - \Delta G^{\circ}_{\text{obs},1}) + (\Delta G^{\circ}_{\text{hyd},2} - \Delta G^{\circ}_{\text{hyd},1}) \quad (\text{eq 1.07})$$

## 1.4 Monitoring Protein Ligand Interactions

One of the most common methods for measuring the affinity of an inhibitor for its targeted receptor is by measuring its half maximal inhibitory concentration ( $IC_{50}$ ). The  $IC_{50}$  is a measure of how much ligand (or inhibitor) is needed to inhibit a given biological process by 50%. Although the  $IC_{50}$  is not a direct measure of the compounds affinity,  $IC_{50}$  and affinity can be related through the Cheng-Prusoff equation, which can be used to calculate  $K_i$  from  $IC_{50}$ .<sup>xvi</sup> For receptor inhibition, the equation is:

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}} \quad (\text{eq 1.08})$$

where  $K_i$  is the equilibrium constant,  $[S]$  is fixed concentration of substrate and  $K_m$  is the concentration of substrate at which enzyme activity is one-half maximal. However, for inhibition constants at cellular receptors, the equation is given by:

$$K_i = \frac{IC_{50}}{1 + \frac{[A]}{EC_{50}}} \quad (\text{eq 1.09})$$

where  $K_i$  is again the equilibrium constant,  $[A]$  is the agonist concentration, and  $EC_{50}$  is the concentration of inhibitor (or agonist) that results in half-maximal activation of the receptor.

In order to obtain the thermodynamic parameters of binding,  $\Delta H^\circ$  and  $\Delta S^\circ$ , one may use the van't Hoff equation to obtain values for  $\Delta H^\circ$  and  $K_a$  at different temperatures, which can then be used to obtain values of  $\Delta G^\circ$  and  $\Delta S^\circ$ . The van't Hoff equation is given by the equation:

$$\frac{d \ln k}{dT} = \frac{\Delta H^\circ}{RT^2} \quad (\text{eq. 1.10})$$

where  $R$  is the gas constant. Following a temperature dependent study of the protein ligand interactions, one can calculate the enthalpic contributions of binding. Alternatively, in the absence of heat capacity changes, one could plot  $\ln k$  versus  $1/T$  such that the slope is  $-\Delta H^\circ/R$ , or using the equation given by:

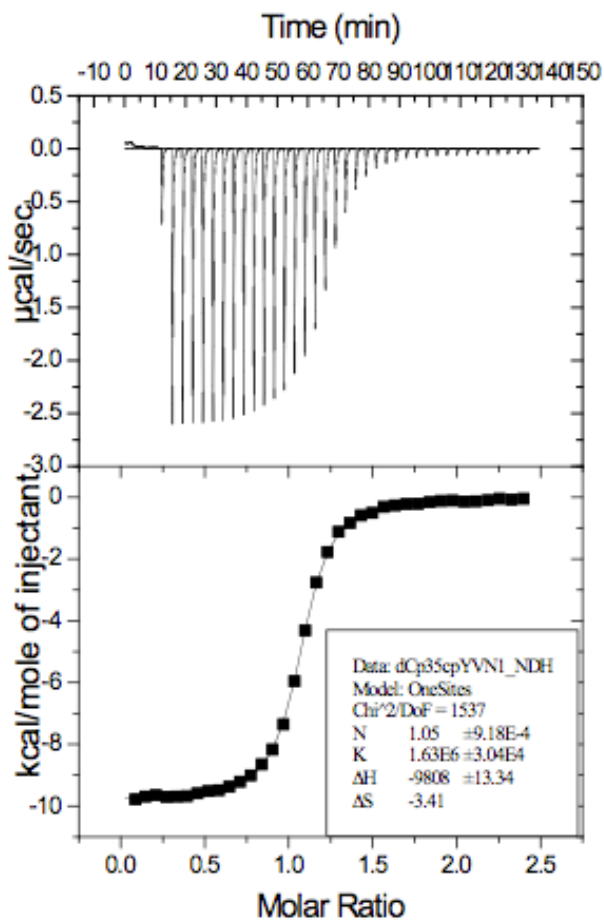
$$\frac{d \ln k}{dT} = \frac{\Delta H^\circ}{R} \quad (\text{eq 1.11})$$

One complicating factor with using the van't Hoff analysis, however, is that it relies on both entropy and enthalpy being temperature independent.<sup>xvii</sup> Because this usually is not the case for protein-ligand interactions, this technique can yield incorrect enthalpy and entropy values that are substantially error prone and do not match those obtained with calorimetric techniques. Although some reports suggested that accurate data can be extrapolated correctly when using the proper techniques and number of trials for the analysis,<sup>xviii,xix,xx</sup> it still remains a challenging method for studying protein-ligand interactions because it requires a large number of experiments relative to other known methods.

Isothermal Titration Calorimetry (ITC) is currently the most used and most reliable analytical technique to measure the incremental changes in heat released or absorbed upon injections of a ligand solution (L) into a solution of a protein (P). ITC

involves a stepwise titration of a protein solution with ligand in a calorimeter so that temperature changes during the interaction are detected directly. Accordingly, when each injection is plotted relative to the ligand concentration in solution, the resulting isotherm of the type shown in **Figure 1.4.1** can be extracted to give each of the thermodynamic parameters of binding.<sup>xxi</sup> Because of the ease and accuracy of extrapolating thermodynamic data, ITC has in recent years become a popular technique for determining the thermodynamic parameters of many ligand-receptor biological processes.<sup>xxii</sup>

**Figure 1.4.1.** Typical ITC thermograph.<sup>xxiii</sup>



Extracting useful data from the titration requires modeling a best-fit curve to the data points, which is typically done using the Wiseman isotherm.<sup>6,xxiv,xxv</sup> For the systems discussed in this chapter, protein (P) and ligand (L) interact in a 1:1 stoichiometry (eqn. 1.03.). Therefore, utilizing the ratio of ligand to receptor, the Wiseman isotherm (eq 1.12) can be employed to determine the stepwise change in heat of the system with respect to moles of ligand added per injection ( $dQ/d[X]_i$ ) relative to the absolute ratio of ligand to receptor ( $X_r = [L]_i/[P]_i$ ):

$$\frac{dQ}{d[X]_t} = \Delta H^\circ V_0 \left[ \frac{1}{2} + \frac{1 - X_r - r}{2((1 + X_r + r)^2 - 4X_r)^{1/2}} \right] \quad (\text{eq 1.12})$$

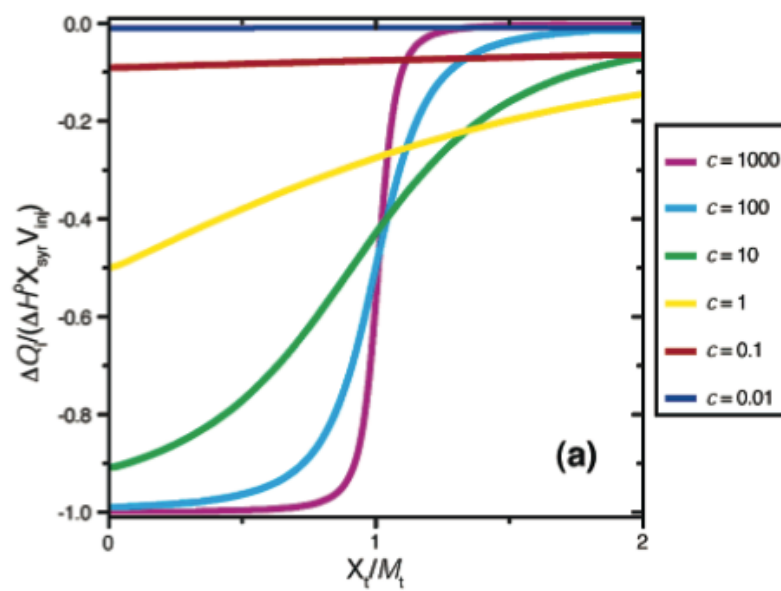
where

$$\frac{1}{r} = c = K_a[P]_t \quad (\text{eq 1.13})$$

and  $V_o$  is the volume of the cell. It was originally noted by Wiseman *et al.*<sup>6</sup> that the shape of the isotherm changes according to the product of the association constant,  $K_a$ , and the concentration of the selected biological receptor, which in this case is the protein, P. This value has been referred to as  $c$ , or the Wiseman coefficient, and it has been suggested that  $c$  values should fall in the range of 10 to 1000 for isotherms that enable the most accurate extraction of thermodynamic data from the titration.<sup>11</sup> The effect that the  $c$  value has on the shape of ITC titrations is shown in Figure 1.4.2. For curves exhibiting a  $c$  value  $< 10$ , it is often difficult to extract an accurate  $\Delta H^\circ$  due to the lack of well defined base and saturation levels. Conversely, when the  $c$  value  $> 1000$ , it is difficult to obtain an accurate fit to the curve due to a lack of data points at the inflection point.<sup>xxvi</sup>



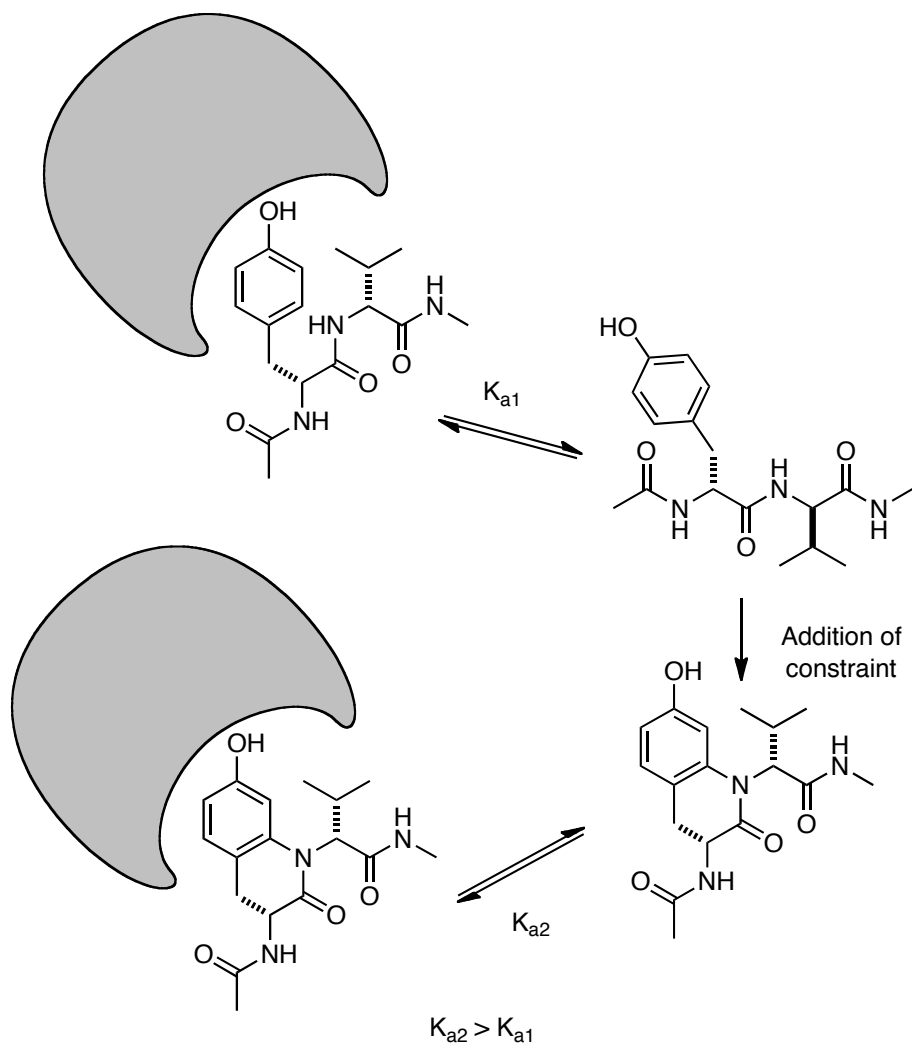
**Figure 1.4.2** Graphs of ITC isotherms showing the differences in  $c$  values.<sup>xxvii</sup>



## 1.5 Ligand Preorganization and Conformational Constraints

Flexible molecules bind to a receptor in a defined conformation to optimally align its functional groups and nonpolar substituents complementarily to those in the binding pocket.<sup>xxviii</sup> The entropic penalty arising from a reduction in the accessible rotameric degrees of freedom of a ligand upon binding to a protein has long been considered avoidable by introducing structural modifications that constrain or decrease the number of rotatable bonds in solution.<sup>xxix</sup> Provided that these modifications do not cause unfavorable interactions within the binding domain, there will theoretically be a smaller entropic penalty upon binding of the constrained analog, which should therefore create favorable gains in the free energy of the binding (shown in Figure 1.5.1).<sup>xxx,xxxi</sup>

**Figure 1.5.1.** Proposed Benefits of Ligand Preorganization

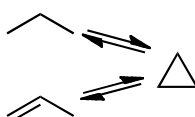

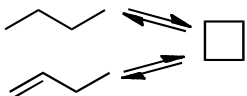
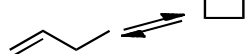
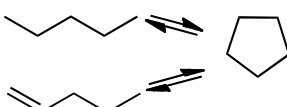



Indeed, when one expands the Gibbs free energy equation to account for changes in the rotational degrees of freedom that are incurred upon binding, as shown in eq 1.14, it can be observed that penalties from restricting rotors have adverse effects on the overall  $\Delta G^\circ$  of binding. Thus, the notion that these penalties can be avoided by constraining the ligand is considered by many to be a viable approach in ligand design.<sup>xxxii</sup>

$$\Delta G^\circ = \Delta G^\circ_{\text{t+r}} + \Delta G^\circ_{\text{rotors}} + \Delta G^\circ_{\text{conf}} + \Delta G^\circ_{\text{h}} + \sum \Delta G^\circ_{\text{p}} + \Delta G^\circ_{\text{vdW}} \quad (\text{eq 1.14})$$

Pioneering studies by Page and Jenks suggested that the entropy lost upon “freezing” a rotor is between 1.1 – 1.5 kcal•mol<sup>-1</sup>.<sup>xxxiii</sup> These values were obtained from analyzing gas phase entropy differences of cycloaliphatic hydrocarbons<sup>xxxiv</sup> between their respective flexible alkane isomers<sup>xxxv,xxxvi</sup> and corresponding constrained alkenes wherein a terminal olefin was used to restrict the rotation across a C-C bond (shown in Figure 1.5.2).<sup>xxxiii</sup> Jenks then applied a correction factor to account for low-frequency motions in the cycloaliphatic hydrocarbons to arrive at  $\Delta S^\circ_{\text{corr.}}$  of 3.7 – 4.9 e.u. per internal rotation. This value has since been supported by subsequent analysis from Williams et. al., who obtained their assessment by studying the differences in the melting points of a series of crystalline alkanes and alkyl carboxylic acids.<sup>xxxvii</sup> These reported values have since garnered considerable attention from the pharmaceutical sciences, since approximating the entropies that may be reclaimed upon restricting ligand rotors in biological systems has significant implications for drug design.<sup>xxxviii,xxxix</sup>

**Figure 1.5.2** Jenks study on rotor restriction.

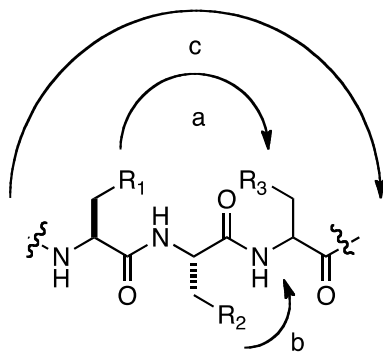
System	- $\Delta S^\circ$ (cal/(molxdeg))	- $\Delta S^\circ$ /no. int. rotors	- $\Delta S^\circ_{\text{corr.}}$ /no. int. rotors
	7.7	3.9	3.9
	7.0	7.0	
<hr/>			
	10.9	3.6	4.9
	10.3	5.1	
<hr/>			
	13.3	3.3	4.8
	13.1	4.4	

Beyond the proposed entropic benefits of ligand preorganization, there are several other suggested reasons to reduce ligand conformational flexibility. First, it is thought

that ligand preorganization may increase bioavailability.<sup>xi</sup> It has been suggested that increasing the rigidity of a peptide or peptide-like molecule may disrupt stabilizing intermolecular interactions that are thought to cause ligand precipitation. Another cooperating benefit of ligand preorganization is a process called peptide mapping, wherein detail of the protein binding pocket is surmised via the synthesis of plausible constrained intermediates when structural data is not easily obtained.<sup>xli</sup> Therefore, since peptides and peptide-related biological inhibitors frequently bind in a rigid conformation such as a  $\beta$ -turn,  $\beta$ -strand, or an  $\alpha$ -helix,<sup>xlii</sup> several strategies exist for preorganizing a ligand to bind in a manner that mimics its bound form.<sup>xliii</sup>

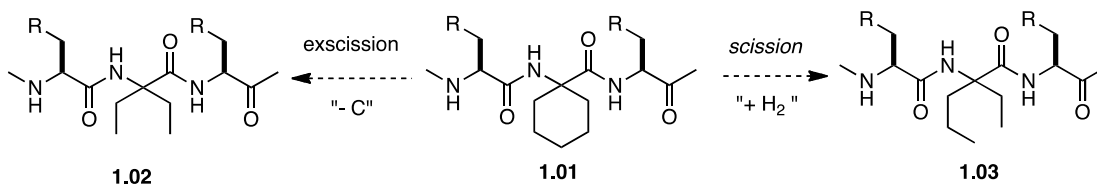
Three main strategies to preorganize a ligand are shown below in Figure 1.5.3.<sup>xliv</sup> The first possibility involves the linking of sidechain  $R_1$  to sidechain  $R_3$ , which is shown as linkage a. Alternatively, the sidechain may be linked to the backbone atoms as shown in path b. Finally, the backbones themselves can be united together forming a macrocycle, as shown by path c.

**Figure 1.5.3.** Modes of peptide preorganization.



To properly determine the effect of ligand preorganization, it is imperative to design and synthesize a flexible control lacking the constraint. We believe it is also critical that the constrained and flexible molecules have the same number and type of heavy atoms, the same functional groups, and the same number of hydrogen-bond donors and acceptors.<sup>xlv</sup> These requirements ensure that differences of activity between the two ligands can best be attributed to the constraint itself. Additionally for macrocyclic ligands, it is a common to employ a scission control (**1.02**), which is the formal hydrogenolysis of a C-C bond, and an excision control (**1.03**), which the deletion of a carbon to provide two linear chains (Figure 1.5.4).<sup>xlix</sup>

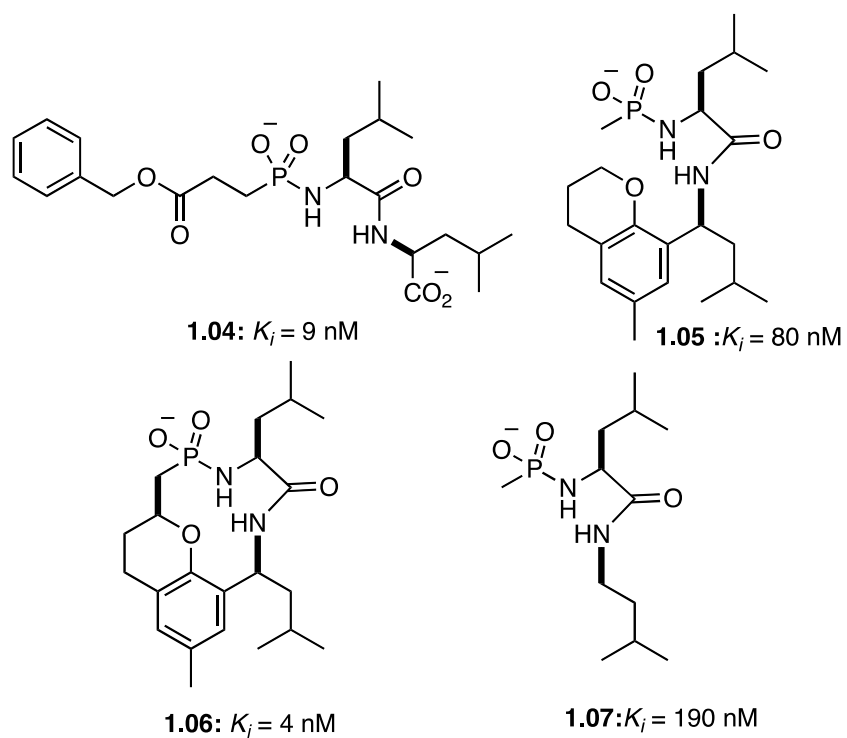
**Figure 1.5.4** Proper scission and excision controls.



## 1.6 Model Systems for Studying Ligand Preorganization

One of the more challenging aspects of designing a ligand with a preorganizational constraint is that one must avoid making unfavorable interactions in the binding pocket that could potentially disrupt key protein residues. In a class of Zinc peptidase thermolysin inhibitors, for example, it was shown through crystallographic data that **1.04** adopts a  $\beta$  turn-like structure in the binding complex (Figure 1.6.1).<sup>xlvi</sup> This finding prompted efforts by Bartlett et. al. to enforce the preferred conformation of **1.04** using a chroman ring as a linker moiety between the C $\alpha$  of the phosphate and the C $\alpha$  of the carboxylate of **1.04**.<sup>xlvii</sup> Accordingly, Ligands **1.05** and **1.06** were designed and synthesized as the respective ligand constraint and flexible control, and the acyclic analogue **1.07** was also made to gauge the effect of the chroman linker itself. Following the synthesis of the desired analogs, the inhibition constants were determined, and these binding studies revealed that macrocycle (**1.06**) bound with higher affinity than both flexible controls. However, it showed only a slight increase in affinity as compared to **1.04**.

**Figure 1.6.1** Bartlett's macrocyclic analogues.

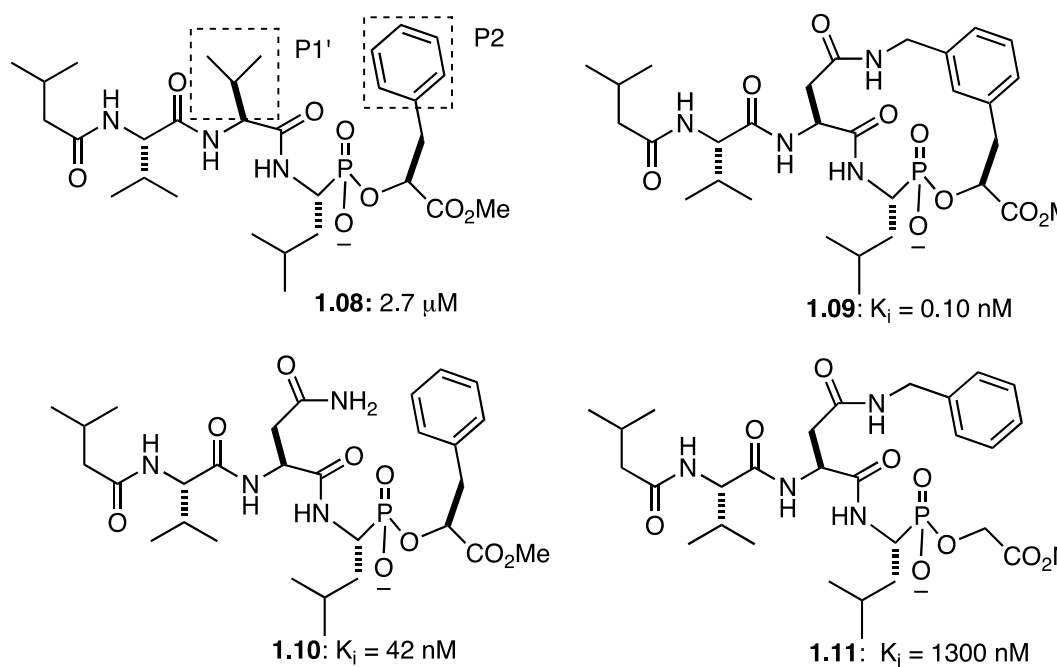




Upon structural investigations into the binding complex of **1.06** and its acyclic controls in the binding pocket, several critical differences in the binding modes became apparent. First, the substitution of the chroman ring for the carboxylate of **1.06** disrupted a critical hydrogen bond that is located in the binding pocket. Second, the chroman ring of **1.05** was rotated by 168° away from its binding pose in **1.06**, which suggests that it is not a proper linker to bridge the two ends of the native inhibitor together. These studies indicate that similar molecules may bind to a receptor in completely different conformations, which makes it difficult to assign the effects of ligand preorganization based solely upon relative binding energetics. Therefore, it is of paramount importance to obtain structural data of both the constrained and flexible ligands to interpret the energetic effects of ligand preorganization as well as to better guide further modifications.

In other experiments by Bartlett, crystal structures of penicillopsin with **1.08** showed a configurational turn in the bound state between the P1' residue and the Y-methyl of the P2 valine (Figure 1.6.2).<sup>xlviii</sup> This finding inspired the investigators to employ structure-based design to synthesize and test the macrocycle **1.09**, which was thought to preorganize the ligand in its bound conformation. Ligands **1.10** and **1.11** were also synthesized as *excision* controls to test for the specific effects of ligand preorganization. Subsequent binding studies revealed that macrocycle **1.09** exhibited an affinity that was 420-fold higher than its acyclic counterpart **1.10** as reflected by  $K_i$ .<sup>xlix</sup>

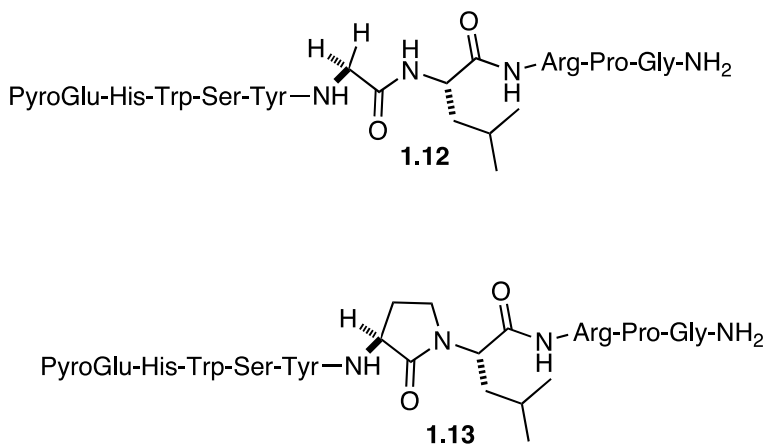
**Figure 1.6.2** Bartlett's pncilopsin-binding ligands.



The results essentially reaffirmed the major tenant of structure-based design wherein the “preorganization” of a ligand into its binding conformation can lead to significant affinity enhancements. However,  $\Delta H^\circ$  and  $\Delta S^\circ$  were not determined for these interactions, so the affinity enhancements cannot be attributed to entropic effects. Nonetheless, they analyzed the differences in rotatable bonds in each of the four ligands to determine that **1.09** had five rotatable bonds in the core structure whereas the linear ligands each had nine. Therefore, both linear ligands bound with different potencies despite having the equivalent number of rotatable bonds, which suggests that there were one or more other factor(s) that had a significant impact on the changes in binding affinity. In subsequent studies to this report, poor correlations between changes in rotatable bonds and binding affinity have also been observed, so there is a question as whether there are absolute, “per rotor” additive effects to the binding entropy that accompany cyclic constraints.<sup>1,ii</sup>

Finally, in a landmark contribution to the field of ligand preorganization in protein-ligand interactions, Freidinger *et al.* targeted a class of luteinizing hormone-releasing receptors (LH-RH). It was proposed that these molecules bound in the  $\beta$ -turn conformation that is shown in Figure 1.6.3.<sup>lii</sup> In an effort to preorganize the ligand into its bound state, they synthesized and tested the lactam derivative **1.13** (Figure 1.6.3) with a bond between a sidechain atom and a backbone atom to force the  $\beta$ -turn. They observed a nearly ten-fold increase in relative potency of **1.12** compared to **1.13**. However, neither the testing of a flexible control nor an investigation into the thermodynamics of binding were pursued and the number of hydrogen bond donors was not held constant.

**Figure 1.6.3** Freidinger study on LH-RH.



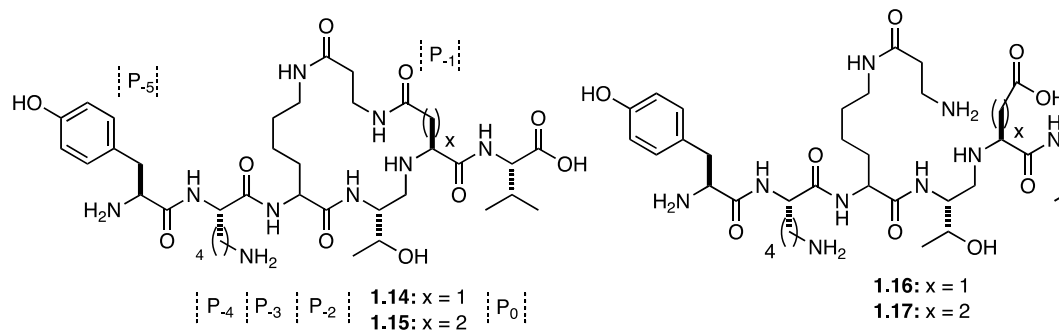
Although these examples support the utility of ligand preorganization in biomolecular interactions, there are many cases wherein the expected affinity boost (1.1 – 1.5 kcal•mol<sup>-1</sup>) per rotor is not obtained.<sup>x,liii,liiv</sup> For instance, Spaller *et. al.* targeted the third PDZ domain (PDZ3) of the postsynaptic density-96 kDa (PSD-95) protein as a model system for designing and testing the utility of preorganizational constraints.<sup>liv</sup> They had additional interest in this system because the PDZ domains are most commonly associated with mediating protein-protein interactions in many eukaryotic cells<sup>lvi,lvii</sup> The

PDZ3 domain was previously shown to bind a nonapeptide such that a four-residue C-terminal stretch made crucial interactions with a  $\beta$ -sheet of the protein domain.<sup>lviii</sup> The Spaller group proposed that a macrocyclic constraint would serve as a viable option to explore the potential utility of ligand preorganization with this system. After modelling studies suggested that the direct linkage of side chains might interfere with the requisite protein–ligand interactions, they examined a series of macrocycles to avoid this problem.

They thus designed constrained ligands **1.14** and **1.15** with varying linker lengths, and the flexible controls **1.16** and **1.17** which were the result of a formal hydrolysis of the amide bond at the  $P_{-1}$  position (Table 1.6.1) were used as controls.<sup>lv</sup> The binding association constants, enthalpy and entropy were determined for this set of ligands utilizing ITC, and the macrocyclic ligands were found to be equipotent to their flexible controls. Interestingly, it was observed for constrained/flexible pair (**1.15/1.17**) where  $x = 2$ , that the constrained macrocycle **1.15** has a more favorable change in entropy than the flexible control **1.17**, which is as expected under ligand preorganization. However, for the ligand pair where  $x = 1$ , **1.14/1.16**, the opposite trend was observed such that the flexible control **1.16** bound with more favorable entropy compared to **1.14**. Unfortunately, no crystal structures of the bound complexes were obtained to better interpret these results. These observations led to two conclusions regarding model systems studying protein ligand interactions. First, it is preferable to design flexible controls that do not introduce polarizable groups such as the  $\text{CO}_2\text{H}$  that are not present in the constraint, as this could result in drastically different binding modes and desolvation energies with the protein. Second, since small changes in the ligand ring size led to sizeable differences in entropy and enthalpy of binding, it was concluded that there are several unknown complexities to

ligand preorganization that are not yet fully understood. Given that the investigators were not able to crystal structures of the complexes

**Table 1.6.1** Bartlett's PDZ-binding macrocycles.

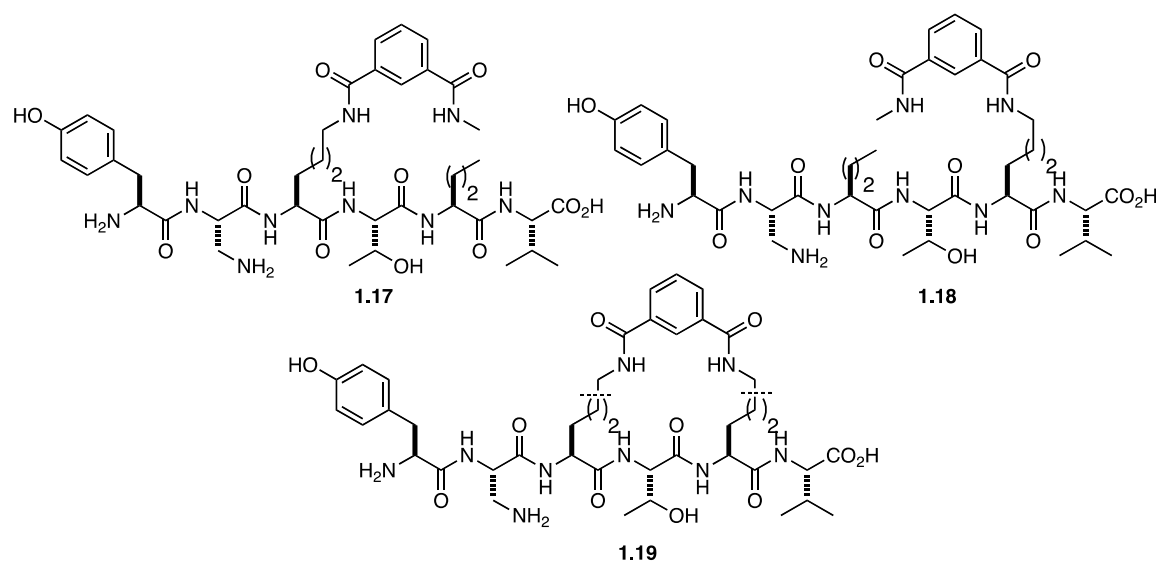


<i>Ligand</i>	$K_a$ ( $\times 10^4$ M)	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>1.14</b>	$2.9 \pm 0.3$	$-6.08 \pm 0.01$	$-2.45 \pm 0.01$	$-3.64 \pm 0.06$
<b>1.15</b>	$2.2 \pm 0.2$	$-5.94 \pm 0.07$	$-1.38 \pm 0.07$	$-4.57 \pm 0.2$
<b>1.16</b>	$4.5 \pm 0.3$	$-7.29 \pm 0.03$	$-2.29 \pm 0.01$	$-5.01 \pm 0.03$
<b>1.17</b>	$5.5 \pm 0.2$	$-7.17 \pm 0.02$	$-4.77 \pm 0.35$	$-2.40 \pm 0.30$

In a follow up to the previous study, the Spaller group envisioned modifying the cyclic constraint of **1.14** and **1.15**. Accordingly, they designed ligands with both benzoate and glutamate bridged macrocycles that are shown in Tables 1.6.1 and 1.6.2, respectively. For the glutamate macrocycle, excision of a methylene group would provide flexible control **1.21** and scission could be utilized to provide controls **1.22–1.24**.<sup>lix</sup> The disassociation constants and binding thermodynamics of these ligands were obtained and are shown in Table 1.6.2. Surprisingly, the constrained mimic **1.19** bound with a slightly higher affinity than the flexible controls **1.17** and **1.18** that was due to a favorable change in *enthalpy*, not *entropy*. For the glutamate-bridged macrocycles, the

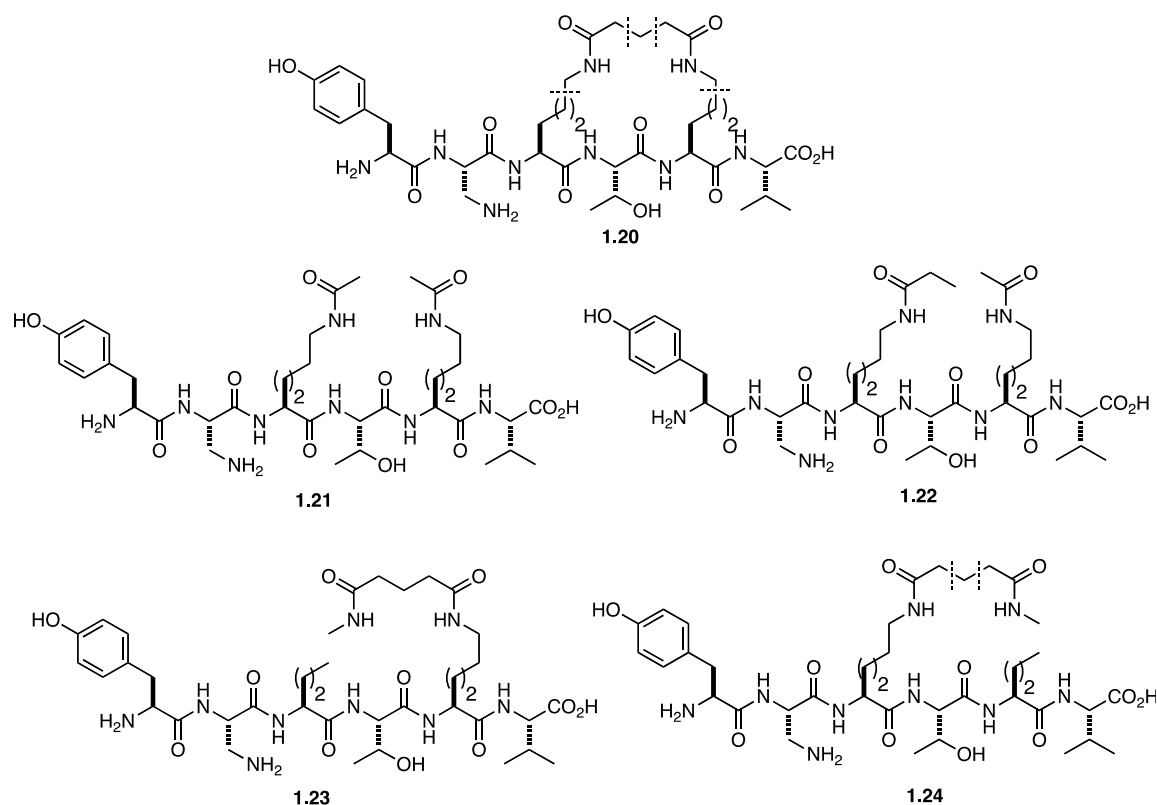
constrained ligand **1.20** bound with nearly equipotent affinities compared to both linear controls. However, the macrocyclic constraint had a more favorable change in *entropy* of binding relative to its linear controls, but this gain was compensated by an unfavorable change in enthalpy.

**Table 1.6.2** Investigation into benzoate-bridged macrocyclic constraints by Spaller.



<i>Ligand</i>	$K_d$ ( $\times 10^5$ M)	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>1.17</b>	2.3 $\pm$ 0.1	-7.3 $\pm$ 0.1	-2.6 $\pm$ 0.1	-4.7 $\pm$ 0.1
<b>1.18</b>	2.1 $\pm$ 1.1	-7.3 $\pm$ 0.1	-3.1 $\pm$ 0.3	-4.2 $\pm$ 0.2
<b>1.19</b>	2.6 $\pm$ 0.1	-7.4 $\pm$ 0.1	-1.9 $\pm$ 0.2	-5.5 $\pm$ 0.1

**Table 1.6.3** Investigation into glutamate-bridged macrocyclic constraints by Spaller



<i>Ligand</i>	$K_d$ ( $\times 10^5$ M)	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>1.20</b>	2.4 ± 0.6	-7.3 ± 0.1	-3.9 ± 0.2	-3.4 ± 0.1
<b>1.21</b>	1.0 ± 0.6	-6.8 ± 0.1	-2.5 ± 0.2	-4.3 ± 0.2
<b>1.22</b>	1.2 ± 0.6	-7.0 ± 0.1	-2.2 ± 0.1	-4.8 ± 0.1
<b>1.23</b>	1.6 ± 0.86	-7.1 ± 0.1	-2.4 ± 0.1	-4.7 ± 0.1
<b>1.24</b>	1.8 ± 0.3	-7.1 ± 0.1	-2.6 ± 0.1	-4.5 ± 0.1

Since the various types of macrocyclic constraints and the corresponding scission and excision flexible controls utilized in this system provided inconsistent results, it was concluded that the supposed benefits of ligand preorganization are not fully understood.

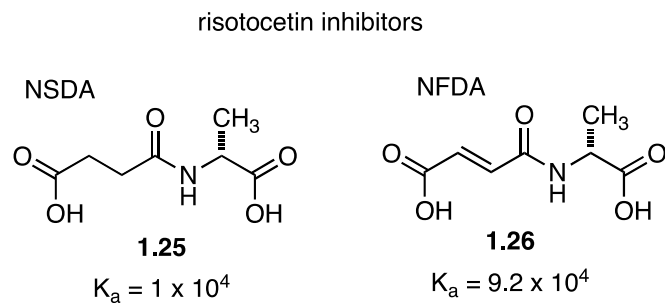


Spaller argues that a significant hindrance to better understanding the effects of ligand constraints is that case studies rarely probe the thermodynamics of binding to determine the enthalpy and entropy of the binding event. Furthermore, systematic studies with structural analysis are not universally performed and the appropriate controls needed to test ligand constraints are frequently ignored.

## **1.7 Olefins as a rigid constraint**

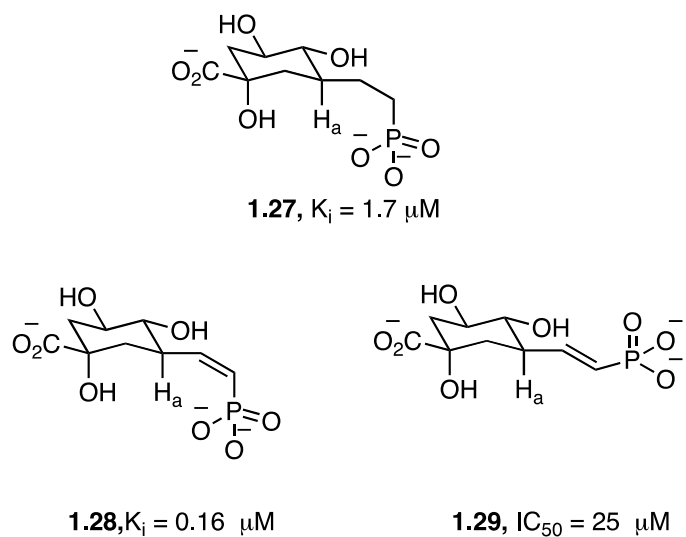
One of the first investigations of using an olefin as a means to restrict rotors was investigated by Williams et al. with a series of risotocetin inhibitors.<sup>xxxii</sup> They found that the alkene containing N-furamyl-D-alanine (NFDA, **1.26**) bound to its receptor with an affinity nearly ten-fold more than that of the flexible ligand N-succinyl-D-alanine (NSDA, **1.25**) (Figure 1.7.1). Calculations suggested that these two ligands bound with no significant difference in buried surface area, so they concluded that a favorable boost in affinity was likely to arise from a favorable entropic component, however, the gains were larger than had been predicted by Jenks et. al.<sup>xxxiii</sup> Thus, it was suggested that there were cooperative benefits that arose from restricting adjacent carbons to the C-C double bond. However, no thermodynamics of binding or structural information was determined for these examples. Thus, there was no direct evidence to correlate the binding affinity with any one thermodynamic parameter.

**Figure 1.7.1.** Williams study of risotcetin inhibitors.



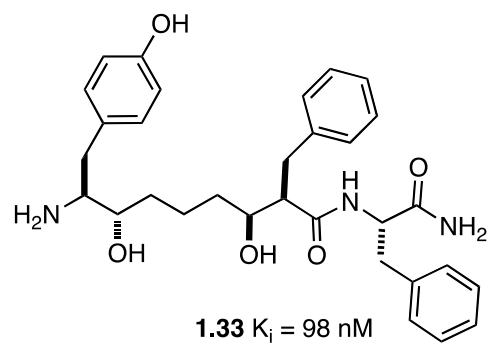
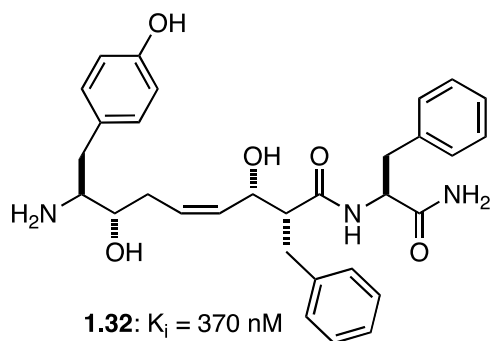
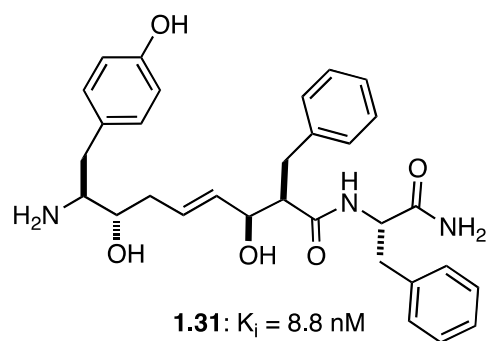
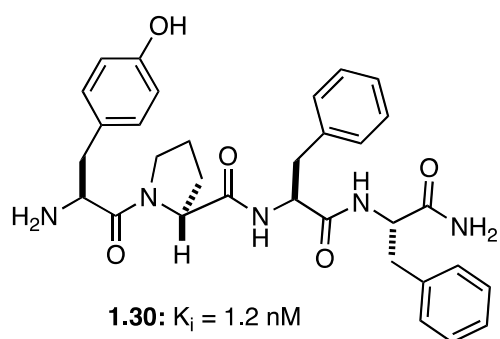
Further contributions by Knowles *et al.* showed that constraining an alkyl linker into its proposed biological conformation with an olefin led to significant increases in binding affinity of several phosphate-derived metabolites of dehydroquinate synthase.<sup>lx</sup> As shown in Figure 1.7.2, the supposed preferred conformation of **1.27** is aligned such that the phosphate group is syn to the atom H<sub>a</sub>. This facilitates the enzymatic deprotonation of H<sub>a</sub> by the enzyme. It was observed that preorganization of the phosphate into such position led to an affinity increase nearly 10-fold that of the parent compound **1.27** and substantially more potent than the corresponding *E* isomer, **1.29**. That **1.28** processed by dehydroquinate synthase provides circumstantial evidence that the internal olefin *does* preorganize the phosphate into its bound conformation. However, thermodynamics of the binding interaction of **1.28** to the enzyme were not determined, so change in entropy or enthalpy of the driving force were not known.

**Figure 1.7.2.** Knowles olefin studies.



Verdine and coworkers have recently reported on one of the most extensive and systematic studies to date regarding the use of an internal olefin as a conformational constraint. Targeting a new class of peptide mimetics for  $\mu$ -opioid receptors, they investigated constraining a ligand into its putative bound conformation utilizing internal olefins as shown in Figure 1.7.3.<sup>lx</sup> Interestingly, they found that there was a strong preference for the *E*-olefin such as **1.31** over the relative *Z*-olefin isomer **1.32**. Furthermore, they showed that the constrained ligand bound better than the flexible control such as **1.33**, but they were not able to achieve any peptide mimetics that bound with greater affinity than the native ligand **1.30**. Although this systematic study supports the utility of olefin constraints, there were no thermodynamic or structural studies to provide any explanation for the observed affinity boosts. Because of the lack of thermodynamic and structural analysis, then the implication of this strategy has no compelling thermodynamic or structural rationale.

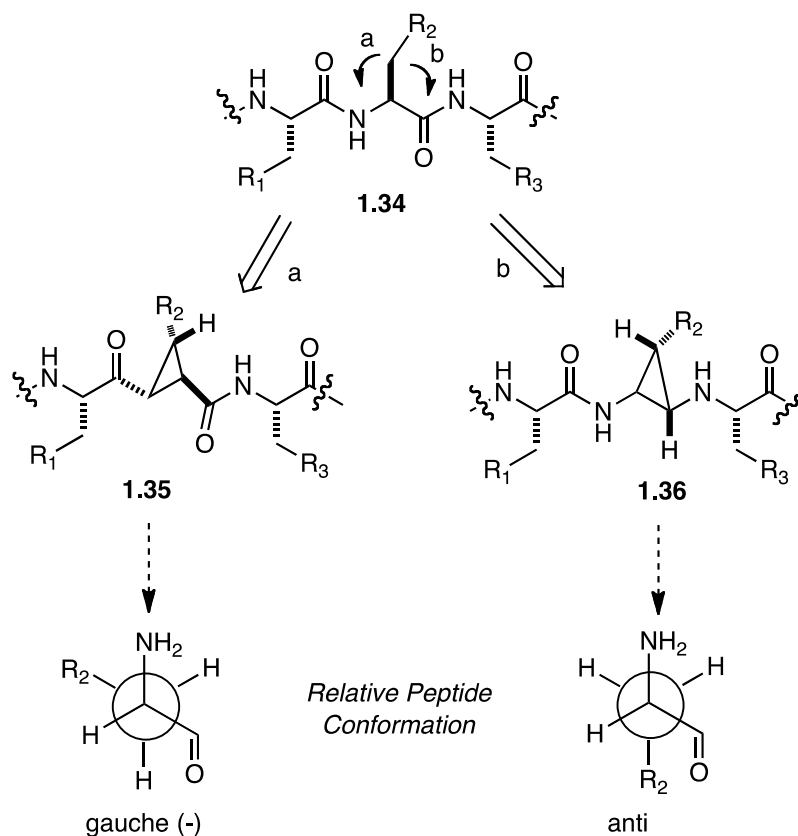
**Figure 1.7.3.** Verdine olefin studies.



## 1.8 Cyclopropyl Constraints

Since peptides serve as leads in the early stages of many drug design endeavors, the Martin group initiated a program nearly 25 years ago to study the effects of ligand preorganization utilizing novel cyclopropane-derived peptide mimetics. Initial molecular modeling studies suggested that 1,2,3-trisubstituted cyclopropanes might serve as rigid and novel replacements for peptides such as **1.34**.<sup>lxii</sup> Peptide mimetic **1.35** was derived from a side chain to backbone constraint (Figure 1.8.1, arrow a) wherein the nitrogen atom of the amide bond was replaced with a carbon atom. Similarly, **1.36** was derived from **1.34** by forming a bond between C ( $\beta$ ) of the side chain and the carbonyl carbon atom of the backbone (arrow b). Modeling studies of the derived ligands then depicted that the *trans* geometry of the cyclopropane in **1.35** not only restricted the conformation of the peptide but also spatially aligned the side chains so that they adopt a  $\chi_1$ -angle of *gauche* (–) ( $-60^\circ$ ) for the corresponding amino acid residue. Similarly, the backbone in **1.36** was aligned so that the side chain was aligned in a  $\chi_1$ -angle of *anti* ( $180^\circ$ ). These model studies on the utility of cyclopropanes paved the road for many future studies on the interaction of conformationally constrained peptides to a host protein/receptor.<sup>lxiii</sup>

**Figure 1.8.1** Early model of cyclopropyl peptide mimetics.

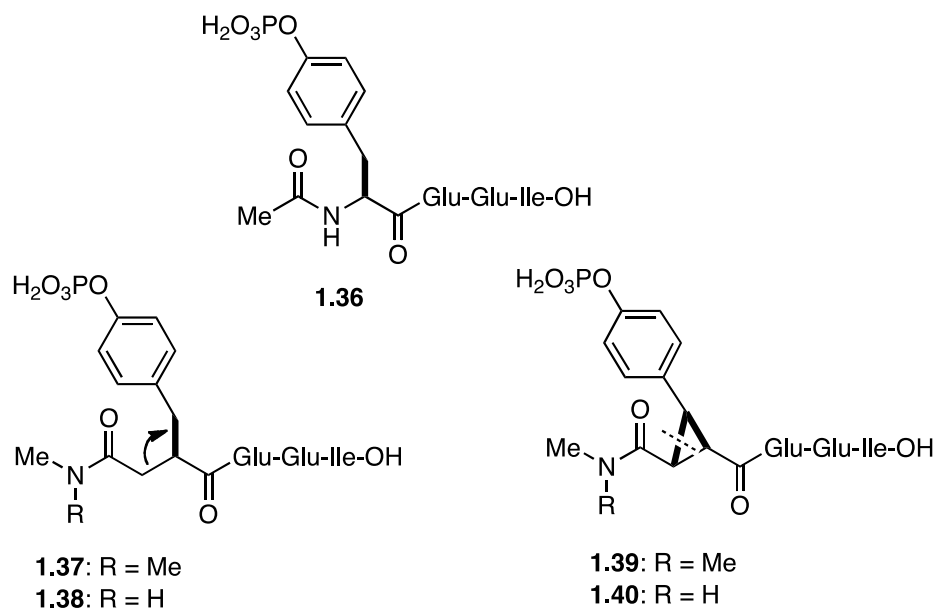


Following a series of cyclopropane-led ventures into a variety of biological receptors, we did not quite find the gains of preorganization that we sought out after.<sup>lxiv, lxv, lxvi</sup> For the most part, the combined efforts of designing and testing ligand constraints in protein-ligand interactions typically led to constrained ligands that were of equipotent binding relative to their flexible counterparts, although there were some notable exceptions.<sup>lxvii</sup>

Given that in many cases binding affinity enhancements did not accompany “supposed” entropic benefits ligand preorganization, we then extended our studies to measure the enthalpies and entropies of the protein-ligand association with (ITC). Early contributions using the Src SH2 domain revealed for the first time that ligand

preorganization in protein-ligand interactions can lead to more favorable entropies of binding compared to a flexible control (Table 1.8.1). Unfortunately, an overall increase in binding affinity was not realized due to a balancing enthalpic.<sup>lxviii, lxix</sup> It was shown through crystallography that the constrained ligand **1.38** adopt a conformation similar to native **1.36** such that the position of their backbone atoms was almost identical. However, crystal structures with the flexible control **1.40** did not provide any useful information.<sup>lxx</sup> Thus unfortunately, the origin of the enthalpic penalty could not be determined from the X-ray data.

**Table 1.8.1.** Src SH2 binding studies.<sup>a</sup>



<i>Ligand</i>	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta H^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$-T\Delta S^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )
<b>1.36</b>	$4.1 \times 10^6$	-9.0	-6.1	-2.9
<b>1.37</b>	$6.3 \times 10^6$	-9.3	-5.9	-4.3
<b>1.38</b>	$1.4 \times 10^7$	-9.7	-6.9	-2.8
<b>1.39</b>	$1.0 \times 10^7$	-9.6	-5.9	-3.7
<b>1.40</b>	$1.7 \times 10^7$	-9.8	-7.3	-2.5

<sup>a</sup>Thermodynamic obtained with Src SH2 domain at 25 °C.

Since crystallographic studies do not provide accurate information regarding the movement and fluctuations in the protein-ligand complex, the structural data that was obtained did not address the possibility that differences in hydrogen bonding in the complex might have contributed to the enthalpic penalty. NMR techniques can reveal insights on the effects of protein fluctuation in molecular associations that are not easily

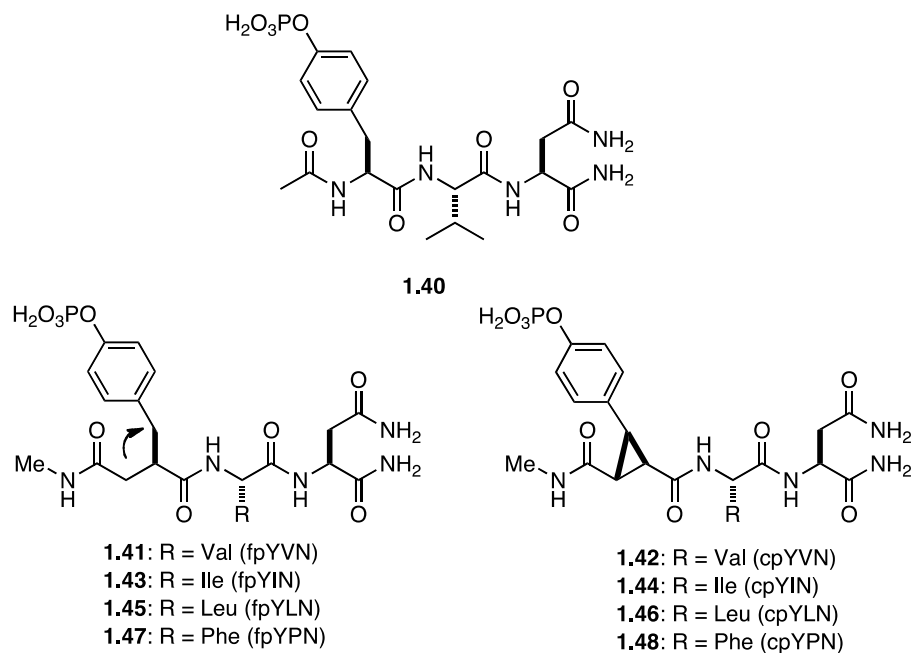


obtained via X-ray.<sup>lxxi, lxxii</sup> One such technique is to measure the chemical shift perturbations (CSPs) of the domain between its bound and unbound form. CSPs of amide bonds can give insight into the hydrogen bond interactions of a protein-ligand complex. CSDs of the pair **1.39** and **1.40** showed that the flexible mimic made stronger hydrogen bond interactions with the domain than did **1.39**. Additionally, the protein-ligand dynamics between the binding of the constrained and flexible mimics and the Src SH2 domain were studied using NMR relaxation. The results revealed that small, but noticeable differences in the hydrogen bonding between the cyclopropane derived mimic compared to the native ligand in the binding pocket were postulated to lead to the unfavorable enthalpic penalty of binding. Coincidentally, these chemical shift differences were so slight that they could not be observed in the structural data. Since the interactions between the protein and the pY+1 ligand residue account for more than 50% of binding free energy,<sup>lxxiii</sup> it was postulated that the use of ligand constraints at binding “hot spots” may have adverse effects to the energetics of binding.<sup>lxxiv</sup>

To gain further insight into the underlying effects of preorganization utilizing a cyclopropane constraint, the project was expanded to include binding of phosphotyrosine-derived pseudopeptides to the Grb2-SH2 domain.<sup>lxxv</sup> A *trans*-cyclopropane containing a tripeptide derivative was found to be slightly more potent than a flexible control derived from a succinate, *but the more favorable binding was due to an enthalpic advantage, not an entropic one* (Table 1.8.2). In fact, the  $\Delta S^\circ$  of binding for the cyclopropane-derived ligand was less favorable than the  $\Delta S^\circ$  of binding of the flexible ligand. This was the first example wherein use of a preorganizational constraint led to an

enhanced binding affinity relative to a flexible control that was driven by a relatively more favorable enthalpy of binding.

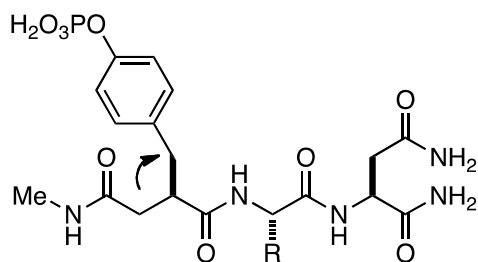
**Table 1.8.2.** Grb2 SH2 binding ligands



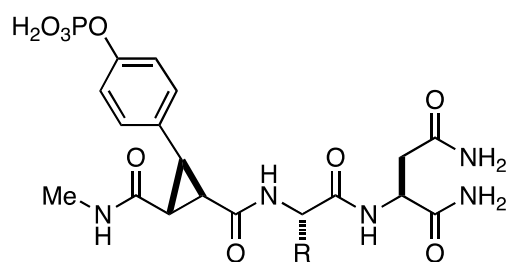
<i>Ligand</i>	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta H^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$-T\Delta S^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )
<b>1.41</b>	$4.5 \times 10^5$	-7.7	-5.4	-2.3
<b>1.42</b>	$2.8 \times 10^6$	-8.8	-7.9	-0.9
<b>1.43</b>	$4.0 \times 10^5$	-7.7	-5.5	-2.2
<b>1.44</b>	$2.1 \times 10^6$	-8.7	-8.3	-0.4
<b>1.45</b>	$1.7 \times 10^5$	-7.1	-4.6	-2.5
<b>1.46</b>	$7.1 \times 10^5$	-8.0	-6.0	-2.0
<b>1.47</b>	$4.1 \times 10^4$	-6.3	-4.2	-2.1
<b>1.48</b>	$2.2 \times 10^5$	-7.3	-5.5	-1.8

The study was then expanded to include constrained and flexible analogues that contained both polar and nonpolar residues at the pTyr+1 position. Those ligands containing the polar residues Gln (**1.49/1.50**), Glu (**1.51/1.52**), and Lys (**1.53/1.54**) generally bound with unfavorable entropies (Table 1.8.3). However, a more favorable enthalpy compensated this overall loss in relative binding *entropy* and the *constrained* ligands bound with a more favorable affinity than the respective flexible controls due to a dominating, more favorable enthalpy change that dominated a less favorable entropy change. Further investigations of this system included studies on proton transfer, temperature dependence, and enthalpy-entropy compensation, these studies did not reveal any significant differences that would provide a reasonable explanation as to the sizeable difference in thermodynamic signatures between the constrained and flexible analogues. Crystal structures of the protein-ligand complex where R = Val (**1.41/1.42**), Ile (**1.43/1.44**), and Gln (**1.49/1.50**) were obtained, and it was shown that constrained ligands consistently made more direct polar contacts in the binding pocket of the complex than the flexible counterparts. Conversely, the flexible ligands made more nondirect, water mediated contacts than the constrained versions.<sup>lxxvi</sup>

**Table 1.8.3. Grb2 SH2 expanded ligands.<sup>a</sup>**



**1.49:** R = Gln (fpYQN)  
**1.51:** R = Glu (fpYEN)  
**1.53:** R = Lys (fpYKN)  
**1.55:** R = Met (fpYMN)



**1.50:** R = Gln (cpYQN)  
**1.52:** R = Glu (cpYEN)  
**1.54:** R = Lys (cpYKN)  
**1.56:** R = Met (cpYMN)

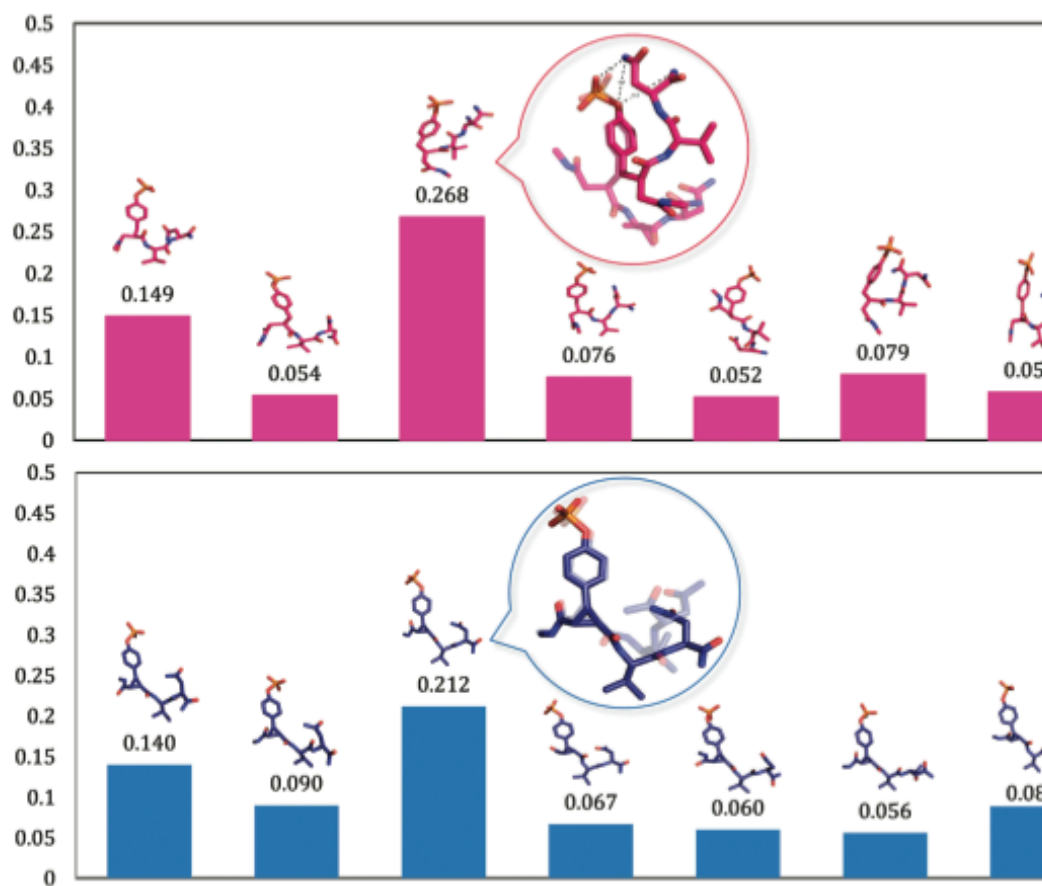
<i>Ligand</i>	$K_a$ (x 10 <sup>5</sup> M <sup>-1</sup> )	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>1.49</b>	5.6±0.7	-7.8±0.08	-8.7±0.8	0.9±0.4
<b>1.50</b>	10.2±0.4	-8.3±0.04	-9.8±0.14	1.5±0.2
<b>1.51</b>	3.0±0.42	-7.5±0.14	-8.8±0.6	1.3±0.7
<b>1.52</b>	36.0±0.4	-7.6±0.06	-10.3±1.12	2.7±0.3
<b>1.53</b>	0.98 ±0.05	-6.8±0.03	-7.7±0.7	0.9±0.2
<b>1.54</b>	5.5±0.3	-7.8±0.04	-9.2±0.6	1.4±0.2
<b>1.55</b>	5.0±0.05	-7.8±0.02	-6.1±0.2	-1.7±0.1
<b>1.56</b>	< 1 X 10 <sup>3</sup>			

<sup>a</sup>Thermodynamic obtained with Grb2 SH2 domain at in HEPES buffer at 25 °C at pH 7.5.

Despite these extensive experiments to systematically test the effects of preorganization in protein-ligand interactions, it was unclear why the flexible ligands bound with more favorable entropy changes than their constrained counterparts. To further probe this scientific question, molecular modeling and computer-generated simulations were undertaken to study the effects of adding constraints to the Grb2 SH2 domain tripeptides. Quasiharmonic analysis correctly predicted that the constrained ligands would bind with more favorable affinities relative to the flexible controls that was driven by a favorable change in enthalpy. Furthermore, upon analyzing the solution conformations that were utilized in the calculations, it was discovered that the flexible ligand **1.41** adopts a more ordered conformation in the solution state compared to the constrained variant (Figure 1.8.2).

The preferred solution structure arises from a stabilizing hydrogen bond between the side chain of the C-terminus and the phosphate group at the pY position.<sup>lxvii</sup> It is worth reiterating that the purpose of introducing constraints in drug design is to *lower* the conformational entropy of the molecule relative to its *more flexible* parent compound, which is the exact opposite conclusion obtained from using computer-generated molecular modeling. Because this discovery calls into great question one of the founding and guiding principles of ligand preorganization, one must ask why the constrained ligand in this study binds with a higher affinity if not for having a more favorable conformation in solution conformations. This study is also one of the first to suggest that lowering the conformational entropy of a ligand in a manner that does not mimic the bound form may still provide more favorable relative entropies compared to a flexible control.

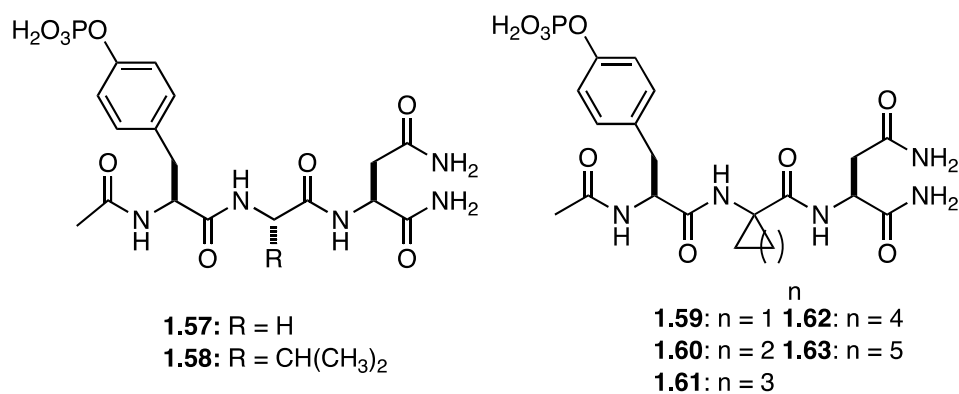
**Figure 1.8.2.** Grouping of population of proposed solution state structures of both constrained and flexible ligands.



## 1.9 Studies of Ligand Binding Affinity Increases Driven by Added Hydrophobicity.

It is generally accepted in drug design that one may increase drug potency by adding nonpolar substituents.<sup>xi</sup> The rationale for this strategy is based upon the hydrophobic effect and the entropically favorable release of water molecules into the surrounding solvent during ligand desolvation.<sup>lxxviii, lxxix</sup> Previous studies in this regard were pursued by Garcia-Echevaira et al. in the search of potent inhibitors of the Grb2 SH2 domain.<sup>lxxx</sup> Since it had been shown that target ligands of the variety pYXN, where X was any nonpolar residue, bound in a  $\beta$ -turn at the pY+1 position (shown in Figure 1.9.1), they initially explored the effects of using  $\alpha,\alpha$ -disubstituted amino acids to preorganize the backbone into its preferred turn.<sup>lxxxi</sup> When the binding affinity of these ligands was analyzed, it was shown that a steady increase in affinity accompanied the increase of methylene carbons up until the six membered analog **1.63**. To obtain some preliminary information as to the cause of this increase, they modeled vdW contacts in the pocket, and these correlated well with the increases in binding affinity. This modeling study suggests that the hydrophobicity of the  $\alpha,\alpha$ -disubstituted residues could be correlated to the incremental increase in binding affinity per methylene unit, but no structural data were obtained to determine enthalpy and entropy of these associations.

**Figure 1.9.1.** Affinities of pY+1 residues.

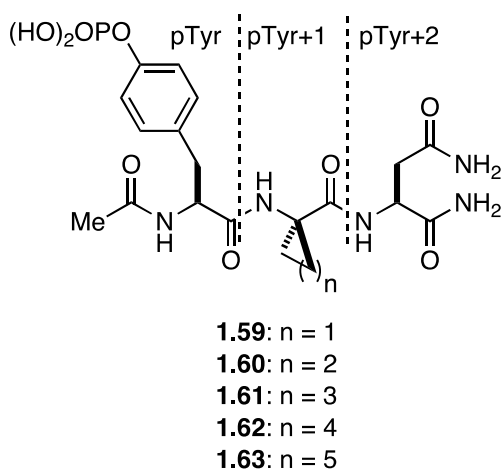


Ligand	IC <sub>50</sub> μM	pTyr <sup>+1</sup> vdW contacts
<b>1.57</b>	67.0	0.0
<b>1.58</b>	4.30	7
<b>1.59</b>	72.5	2
<b>1.60</b>	18.2	2
<b>1.61</b>	7.90	4
<b>1.62</b>	0.21	8
<b>1.63</b>	1.11	8



The Martin group recently reported on the thermodynamics of binding between a series of  $\alpha,\alpha$ -cycloaliphatic peptides shown in Figure 1.9.1 and the Grb2 SH2 domain.<sup>lxxxii,10</sup> The ligands were synthesized and their complexes with the Grb2 SH2 domain was studied with ITC to determine the  $\Delta S^\circ$  and  $\Delta H^\circ$  of binding (Table 1.9.1). For the pTyr+1 substituted ligands where  $n = 1 - 5$ , it was found that the binding affinities of ligands **1.59-1.62** increased incrementally with ring size up to a maximal affinity where  $n = 4$ , which was nearly equipotent to the cycloheptyl analog ( $n = 5$ ). ITC studies revealed that the binding affinity was driven by relatively more favorable changes in *enthalpies*, not *entropies*, upon the addition of methylene groups which is counter to what would be predicted based off of conventional wisdom regarding the thermodynamic effects of adding apolar surface area to an inhibitor.<sup>xi</sup> X-ray crystallographic (X-ray) analysis of these ligands complexed with the Grb2-SH2 domain were obtained, and these revealed that an incremental increase in van der Waals (vdW) contacts made in the binding pocket correlated very well to the favorable changes in  $\Delta H^\circ$  that were observed per addition of each methylene group.<sup>lxxxiii</sup> However, the incremental changes in enthalpy did not correlate with  $\Delta C_p$  that was determined for each ligand. This is unexpected considering that the temperature-dependent studies of protein-ligand interactions are commonly thought to be a barometer of buried apolar surface area. Given the unexpected nature of the trends in enthalpy that were obtained with this study along with the failings of  $\Delta C_p$  to reflect the increase in apolar surface area, the authors concluded that there many misunderstood complexities to correlating structural changes with thermodynamics in protein-ligand interactions.

**Table 1.9.1** pY+1 ligand study that bound to the Grb2 SH2 domain.

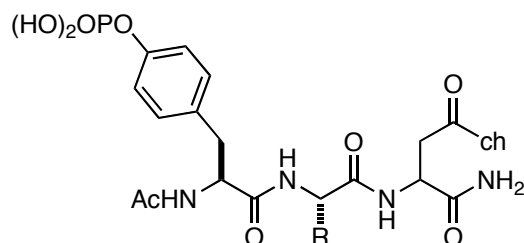


<i>Ligand</i>	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )	pTyr+1 vdW contacts	$\Delta C_p$ (cal•mol <sup>-1</sup> K <sup>-1</sup> )
<b>1.59</b>	<b>-7.1 ± 0.1</b>	<b>-3.3 ± 0.3</b>	<b>-3.8 ± 0.1</b>	<b>4</b>	<b>-116 ± 12</b>
<b>1.60</b>	<b>-7.7 ± 0.1</b>	<b>-5.4 ± 0.3</b>	<b>-2.3 ± 0.2</b>	<b>5</b>	<b>-185 ± 8</b>
<b>1.61</b>	<b>-8.5 ± 0.1</b>	<b>-6.3 ± 0.4</b>	<b>-2.2 ± 0.2</b>	<b>9</b>	<b>-141 ± 8</b>
<b>1.62</b>	<b>-9.3 ± 0.1</b>	<b>-8.5 ± 0.4</b>	<b>-0.8 ± 0.4</b>	<b>13</b>	<b>-181 ± 10</b>
<b>1.63</b>	<b>-8.9 ± 0.1</b>	<b>-6.8 ± 0.3</b>	<b>-2.1 ± 0.2</b>	<b>14</b>	<b>-173 ± 8</b>

The Martin group extended their investigations of hydrophobic effects by evaluating Grb2-SH2 binding ligands with linear alkyl substituents at the pY+1 residue. The ligands shown in Table 1.9.2 were synthesized and their binding parameters were obtained by ITC.<sup>lxxxiv</sup> Although a favorable increase in binding free energy were observed upon adding methylene groups to Ala, subsequent addition of apolar surface area did not

provide any significant increases in binding affinity. Any gains that were obtained by a more favorable change in enthalpy were compensated by a loss in entropy. It was also observed that an increase in vdW contacts made in the binding pocket correlated with slight gains in enthalpy upon the addition of apolar surface area, but these parameters taken together showed a poor correlation with  $\Delta C_p$ .<sup>o</sup> Interestingly structural studies of the bound complexes showed that when R = Nva, the alkyl chain adopts a gauche conformation in the bound complex which is also observed in ligands **1.61**, **1.62** and **1.63**. Thus, the plateau in observed binding affinity could be arising from the unfavorable interaction that the alkyl chains make in the pocket.

**Table 1.9.2** Binding studies of linear pY+1 ligands with the Grb2-SH2 domain in HEPES at 298 K.

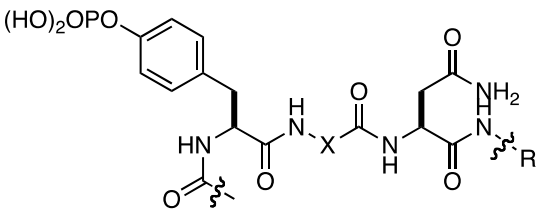


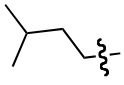
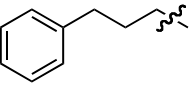
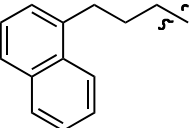
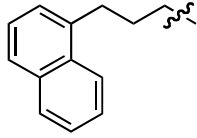
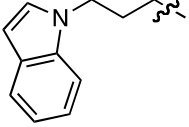
Compound	R	Xaa
<b>1.64</b>	Methyl	Ala
<b>1.65</b>	Ethyl	Abu
<b>1.66</b>	<i>n</i> -Propyl	Nva
<b>1.67</b>	<i>n</i> -Butyl	Nle
<b>1.68</b>	<i>n</i> -Pentyl	Hnl

<i>Ligand</i>	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )	pTyr+1 vdW contacts	$\Delta C_p$ (cal•mol <sup>-1</sup> K <sup>-1</sup> )
<b>1.64</b>	<b>-7.3 ± 0.1</b>	<b>-5.1 ± 0.1</b>	<b>-2.2 ± 0.4</b>	<b>4</b>	<b>-123 ± 9</b>
<b>1.65</b>	<b>-8.1 ± 0.1</b>	<b>-6.8 ± 0.5</b>	<b>-1.3 ± 0.1</b>	<b>9</b>	<b>-170 ± 15</b>
<b>1.66</b>	<b>-8.0 ± 0.1</b>	<b>-6.7 ± 0.5</b>	<b>-1.3 ± 0.6</b>	<b>12</b>	<b>-173 ± 13</b>
<b>1.67</b>	<b>-8.1 ± 0.1</b>	<b>-7.3 ± 0.3</b>	<b>-0.8 ± 0.2</b>	<b>14</b>	<b>-138 ± 12</b>
<b>1.68</b>	<b>-8.0 ± 0.1</b>	<b>-7.2 ± 0.3</b>	<b>-0.9 ± 0.2</b>	<b>14</b>	<b>-148 ± 7</b>

Garcia-Echevaria et al. probed the effects of adding nonpolar surface area to the pY+3 residue of tripeptides known to bind to the Grb2 SH2 domain.<sup>lxxxv,lxxxvi</sup> They reasoned that adding such nonpolar substituents at that position would make favorable interactions with an extended hydrophobic patch of the Grb2 SH2 domain. Indeed, they observed a significant increase in binding affinity when nonpolar substituents were added (Figure 1.9.2). Modeling programs suggested that these gains in ligand binding affinity were due to favorable overlap between the apolar R groups and an extended nonpolar patch of the binding pocket, but no structural or thermodynamic data was determined to confirm this hypothesis.

**Figure 1.9.2.** pY+3 derived ligands.



compd	R =	X =	IC <sub>50</sub> (μM)
1.69		Ile	3.2 ± 0.1
1.70		Ile	2.3 ± 0.1
1.71		Ile	0.33 ± 0.1
1.72		Ac <sub>6</sub> C	0.047 ± 0.1
1.73		Ac <sub>6</sub> C	0.0012 ± 0.1
1.74	H <sup>+</sup>	Ac <sub>6</sub> C	0.21 ± 0.1

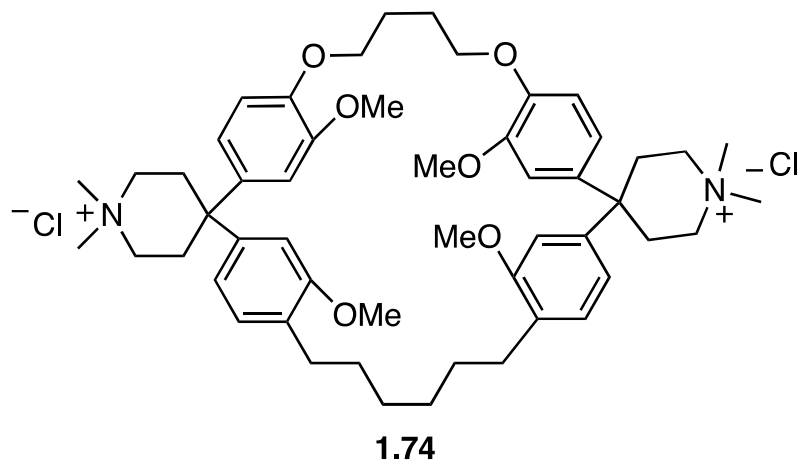
## 1.10 Studies on the Thermodynamic Nature of Hydrophobic Effects

Although the Garcia-Echevaria work clearly displays the potential utility of exploiting the use of added ligand hydrophobicity at the pY+3 position to achieve gains in binding affinity, the fundamental processes that drive such hydrophobic interactions are still poorly understood. There are very few cases that study such “entropy-oriented” modifications in which entropy and enthalpy of binding are determined, and even fewer that correlate the binding thermodynamics with structural changes in the bound complex. To address this pressing need, Ladbury and coworkers have developed a database named SCORPIO (short for Structure/Calorimetry of Reported Protein Interactions Online) that compiles published thermodynamic binding studies for which structural data is also available.<sup>lxxxvii</sup> Upon searching this database for correlations between binding thermodynamics and buried apolar ligand surface area, they found a modest correlation ( $R^2 = 0.70$ ) of buried apolar surface area with Gibbs free energy. This study suggests that while the increase of ligand hydrophobicity may eventuate into an increase in binding affinity, the driving force behind such affinity gains is poorly understood. As such, it further brings into question the prevailing thought that such interactions are entropy driven.

There is also a growing body of work showing that hydrophobic effects may also have strong enthalpic component.<sup>lxxxviii</sup> It was noted by Diederich et al. that several benzene derivatives bound to a host cyclophane **1.74** with extraordinarily large association constants. When these binding parameters were decomposed to give the enthalpy and entropy using a van't Hoff analysis, it was observed that the binding event was strongly enthalpy driven. Furthermore, the binding interactions were severely

weakend when the experiment was conducted in CD<sub>3</sub>OD, yet they still observed that the geometry of the complexes in either water or methanol were conserved. Thus, the investigators concluded that the favorable enthalpies of binding were largely due to the role of water; a finding that was counterintuitive to the generally accepted model of hydrophobic effects (Figure 1.10.1).<sup>lxxxix</sup> This result gave rise to the term “nonclassical” hydrophobic effect to describe host-receptor interactions in water that are enthalpy, not entropy driven.

**Figure 1.10.1.** Diederich hydrophobicity studies.



Ligand	$\Delta G^\circ$ (kcal/mol) in D <sub>2</sub> O, [CD <sub>3</sub> OD]	$\Delta H^\circ$ (kcal/mol) in D <sub>2</sub> O [CD <sub>3</sub> OD]	$T\Delta S^\circ$ (kcal/mol) in D <sub>2</sub> O [CD <sub>3</sub> OD]
<i>p</i> -dicyanobenzene	-5.2 [-1.9]	-9.5 [-4.2]	-4.8 [2.4]
<i>p</i> -dimethoxybenzene	-5.4 [-1.2]	-10.2 [-4.4]	-4.3 [3.2]

Indeed, several thermodynamic studies of biological associations between small molecules and a host protein have also shown that nonpolar ligands can bind in an enthalpy driven process. For example, Ross discovered that small organic compounds



such as indole or proflavin (see Figure 1.10.2) bind to their respective host protein with a *favorable enthalpy* and an *unfavorable entropy* upon binding.<sup>xc</sup> Although this result is the opposite of what may be expected from hydrophobic effects, he suggests that the enthalpy of these interactions stems from stabilizing intermolecular interactions between the ligand and receptor that immediately follow ligand desolvation. Ross, citing corroborating analyses that support this claim,<sup>xc<sub>i</sub>,xc<sub>ii</sub></sup> then postulated that *enthalpy*-driven associations between a host and a receptor represent the norm in biomolecular interactions rather than the exception.

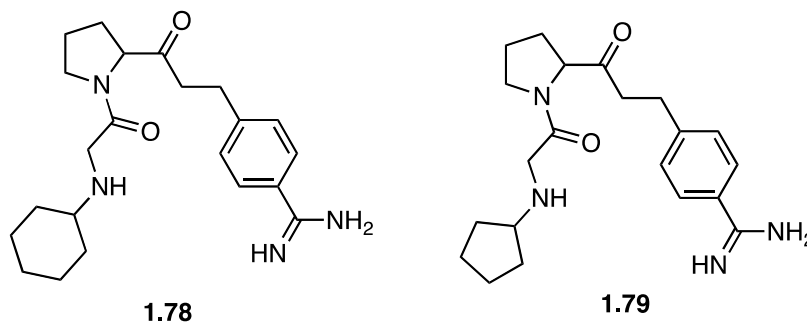
**Figure 1.10.2.** Ross ligand analysis.

Ligand	$\Delta G^\circ$ (kJ/mol)	$\Delta H^\circ$ (kJ/mol)	$\Delta S^\circ$ (J/mol•K)
Lactic Dehydrogenase			
NAD <sup>+</sup>	-3.8	-6.3	-8.3
NADH	-7.4	-6.9	1.7
ADP-ribose	-4.6	-7.6	-10.0
iodosalicylic acid	-3.5	-21.7	-61.0
Chymotrypsin			
proflavin	-6.0	-11	-18
indole	-4.3	-15	-37
N-acetyl-D-Trp	-3.3	-19	-53
hydrocinnamic acid	-2.6	-26	-80

Klebe and coworkers have recently studied the effects in the addition of methylene groups to cyclic residues in a series of thrombin inhibitors.<sup>xc<sub>iii</sub></sup> A key finding was that changing an aliphatic substituent from cyclopentyl in **1.79** to cyclohexyl in **1.78** (**Figure 1.10.3**) resulted in a much more favorable increase in  $\Delta S^\circ$  that was compensated by an unfavorable change in  $\Delta H^\circ$ , and thus the overall binding affinity was unaffected by

the change in ligand hydrophobicity. That the increase in ligand hydrophobicity led to a gain in favorable entropy was interesting, and this prompted structural studies. Interestingly, these studies showed that the change in the binding thermodynamics seen in **1.79** and **1.78** to the increased flexibility of the cyclohexyl side chain that presumably gives an entropic advantage upon binding. Thus, this study demonstrates the importance of correlating thermodynamic studies with structural observations to better understand the binding interactions.

**Figure 1.10.3.** Klebe ligands.

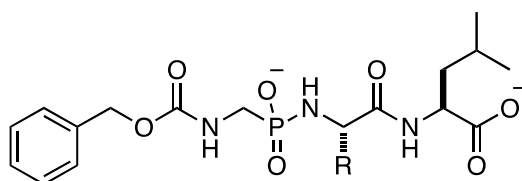


<i>Ligand</i>	$K_a$ ( $\times 10^5 \text{ M}^{-1}$ )	$\Delta G^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta H^\circ$ ( $\text{kcal mol}^{-1}$ )	$-T\Delta S^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )
<b>1.78</b>	$10.2 \pm 0.4$	$-8.3 \pm 0.04$	$-2.3 \pm 0.14$	$6.0 \pm 0.3$
<b>1.79</b>	$1.5 \times 10^6$	$-8.4 \pm 0.04$	$-4.4 \pm 0.14$	$4.0 \pm 0.3$

In further studies to probe hydrophobic effects in protein-ligand interactions, Klebe investigated a class of thermolysin inhibitors.<sup>xciv</sup> Thermolysin belongs to a family of enzymes that can be classified as the M4 family of metalloproteases which can be found in bacteria and fungi. It is considered a prototypical model system for studying protein-ligand interactions as it is robust and has known procedures to provide readily

accessible crystallographic data,<sup>xcv</sup> and as such, it has been used as model system for other pharmaceutically relevant proteases such as the angiotensin converting enzyme (ACE).<sup>x cvi</sup> Klebe explored the effects of adding branched alkyl chains to phosphoramidite-based inhibitors of thermolysin shown in Table 1.10.1. In most cases, they observed a favorable increase in binding affinity that was driven by a dominating enthalpy upon the addition of apolar surface area. This trend was attributed to a relatively *dehydrated* thermolysin binding pocket<sup>x cvii</sup> that was able to make favorable dispersive interactions with the branched alkyl substituents.

**Table 1.10.1** Klebe ligand with modified alkyl surface area.



**1.80:** R = Me

**1.81:** R = *i*Pr

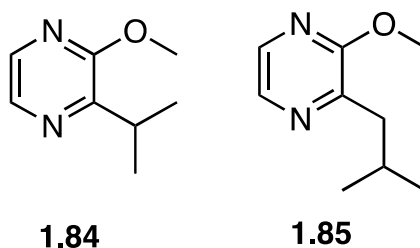
**1.82:** R = *i*Bt

**1.83:** R = Bn

Modification of R	$\Delta\Delta G_{kd}^{\circ}$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H^{\circ}$ (kcal •mol <sup>-1</sup> )	$-\Delta T\Delta S^{\circ}$ (kcal•mol <sup>-1</sup> )
<b><i>i</i>Pr → <i>i</i>Bu</b>	0.9	1.0	-0.1
<b><i>i</i>Bu → Bn</b>	-0.3	1.5	-1.8
<b><i>i</i>Pr → Bn</b>	0.5	2.5	-2.0

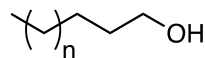
Further reports of enthalpically driven hydrophobic effects have come from the Homans lab in their studies of the incremental addition of methylene groups to pyrazines that bind to mouse Major Urinary Protein I (MUP-I). The Homans group studied the differences in binding thermodynamics of 2-methoxy-3isopropylpyrazine (**1.84**) and 2-

methoxy-3isobutylpyrazine (**1.85**), which contains one additional methylene group, with MUP-I.<sup>xcviii</sup> Remarkably, they found that the binding affinity increased from **1.84** to **1.85** upon the addition of a methylene group, but that this increase in affinity was attributed to a more favorable change in the *enthalpy* of binding relative to **1.84** (Table 1.10.2.). Structural studies showed that there were no bound water molecules in the complexes of **1.84** and **1.85**, which is not unusual as MUP-I is thought to have an unoptimally hydrated binding cavity.<sup>xcviii</sup> Furthermore, the Stone group also reported on a series of thiazole heterocycles that also bound to MUP-I such that favorable changes in enthalpy were observed upon the addition of methylene groups.<sup>xcix</sup> It was noted by both Stone and Homans that the observation that addition of methylene groups led to favorable changes in enthalpy, not entropy, was perplexing because the burial of apolar surface area was considered to grant favorable changes in binding entropy. It was later *proposed*<sup>xii</sup> that since the binding cavity of MUP-I is “dewetted”, then protein-ligand dispersive interactions made in the pocket would not be completely balanced by a corresponding loss in protein-solvent interactions prior to binding<sup>cxx</sup>. Thus, this phenomenon could provide an explanation for the favorable changes in enthalpy observed in both the Stone and Homans studies.

**Table 1.10.2** Homans study of increasing hydrophobicity of pyrazine derivatives

	$\Delta G^\circ$	$\Delta H^\circ$	$-T\Delta S^\circ$
<i>Ligand</i>	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )
<b>1.84</b>	$-8.1 \pm 0.02$	$-10.6 \pm 0.02$	$2.5 \pm 0.8$
<b>1.85</b>	$-9.2 \pm 0.02$	$-11.5 \pm 0.02$	$2.2 \pm 0.1$

To further study the effects of adding hydrophobicity to MUP-I binding ligands, the Homans group extended their study of ligand hydrophobicity to alkyl alkanols. Since they previously postulated that the imbalance in dispersive interactions created by *dehydrated* receptors<sup>xcvii</sup> such as MUP-I could account for enthalpically driven binding interactions, they suggested that the addition of methylene carbons to hexanol would be an ideal model system to test this hypothesis. Accordingly, they determined the thermodynamics of binding between small hydrophobic alkanols (Table 1.10.3) and MUP-I, which showed that the addition of nonpolar surface area was indeed accompanied by *enthalpically* driven gains in affinity.<sup>xiv</sup> These results further supported the hypothesis that the addition of apolar surface area to a protein binding ligand may not always lead to more favorable changes in binding entropy, but that enthalpy changes can also play a major role.

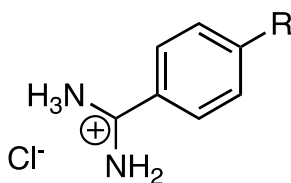
**Table 1.10.4** Homans ligand analysis with MUP-I utilizing alkanols.

<i>Ligand</i>	<i>n</i>	$\Delta G^\circ$	$\Delta H^\circ$	$T\Delta S^\circ$
		(kcal/mol)	(kcal/mol)	(kcal/mol)
<b>1.86</b>	<b>1</b>	$-5.5 \pm 0.02$	$-9.8 \pm 0.02$	$-4.3 \pm 0.8$
<b>1.87</b>	<b>2</b>	$-6.8 \pm 0.02$	$-11.4 \pm 0.02$	$-4.6 \pm 0.1$
<b>1.88</b>	<b>3</b>	$-7.7 \pm 0.01$	$-12.8 \pm 0.01$	$-5.0 \pm 0.1$
<b>1.89</b>	<b>4</b>	$-8.5 \pm 0.02$	$-13.9 \pm 0.02$	$-5.4 \pm 0.1$
<b>1.90</b>	<b>5</b>	$-9.3 \pm 0.05$	$-15.2 \pm 0.05$	$-5.9$

The Homans studies stand in distinct contrast to investigations by Engberts on the addition of methylene groups to Trypsin binders.<sup>c</sup> They studied the effect of adding hydrophobic groups to a set of benzamidine ligands and determined their binding thermodynamics using ITC (Table 1.10.4) It was observed that the addition of methylene groups from **1.91** (R = H) to **1.97** (R = nHex) was accompanied by favorable change in binding *entropy* but an unfavorable change in binding *enthalpy*, such that affinities changes were minimal across the series. They also reported a strong correlation to  $\Delta C_p$ , which served as an indicator of the burial of apolar surface area in protein-ligand binding. They then why they were not observing significant increase in binding affinity that accompanied the burial of hydrophobic surface area. They then conducted binding simulations which suggested that there was no significant steric interaction made in the

pocket that could account for the compensating unfavorable change in enthalpy. The investigators suggested that there are still unknown factors in protein-ligand interactions that make it difficult to achieve marked binding affinity increases due to the modification of a ligands hydrophobic surface area.

**Table 1.10.5** Enberts study of increasing methylenes to Trypsin binders.

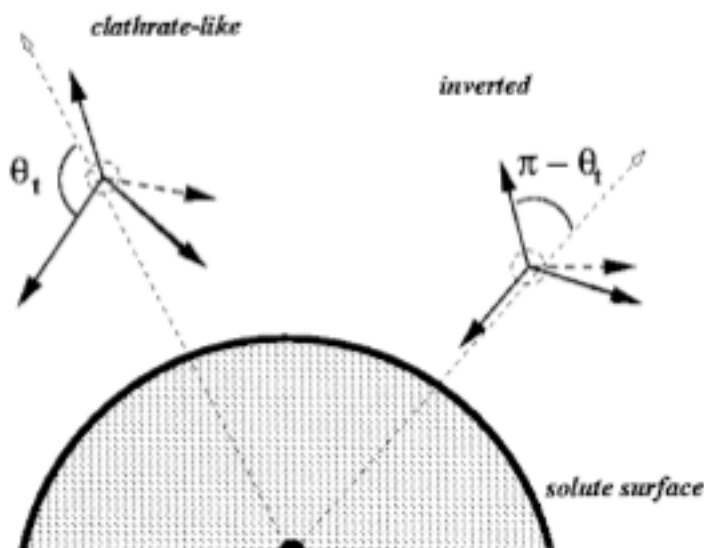


Ligand	$K_a$ ( $\times 10^4 \text{ M}^{-1}$ )	$\Delta G^\circ$ (kcal/mol)	$\Delta H^\circ$ (kcal/mol)	$-T\Delta S^\circ$ (kcal/mol)	$\Delta C_p$ (cal $\cdot$ mol $^{-1}$ K $^{-1}$ )
<b>1.91 (R = H)</b>	<b>4.5 <math>\pm</math> 0.2</b>	-6.4	-4.5	-1.8	-96 $\pm$ 5
<b>1.92 (R = Me)</b>	<b>6.9 <math>\pm</math> 0.3</b>	-6.6	-4.4	-2.2	-100 $\pm$ 3
<b>1.93 (R = Et)</b>	<b>2.9 <math>\pm</math> 0.3</b>	-6.1	-3.3	-2.8	-166 $\pm$ 3
<b>1.94 (R = <i>n</i>Pr)</b>	<b>3.1 <math>\pm</math> 0.2</b>	-6.1	-3.0	-3.1	-143 $\pm$ 0.5
<b>1.95 (R = <i>n</i>Bu)</b>	<b>3.8 <math>\pm</math> 0.2</b>	-6.3	-2.4	-3.9	-174 $\pm$ 1
<b>1.96 (R = <i>n</i>Pent)</b>	<b>5.8 <math>\pm</math> 0.3</b>	6.5	-2.4	-4.1	-151 $\pm$ 2
<b>1.97 (R = <i>n</i>Hex)</b>	<b>13 <math>\pm</math> 0.6</b>	7.0	-2.5	-4.5	-203 $\pm$ 2



It is also worth noting that recent studies question the hypothesis that all water molecules surround apolar surface area in a well defined and highly-ordered manner.<sup>ci</sup> During the studies of the hydrophobic interactions between membranes and their surrounding proteins, it was discovered that there was a large structural dependence on the topology of apolar surface that influenced the hydrogen bond networks formed between water molecules. As shown in Figure 1.10.4, a convex surface led to optimal hydrogen bonding (shown as vectors) between solvent, whereas a flat or concave surface leads to unoptimized hydrogen-bonding networks. The significance of this study is that it suggests that water networks are dependent on the classification of ligand surfaces and topology.

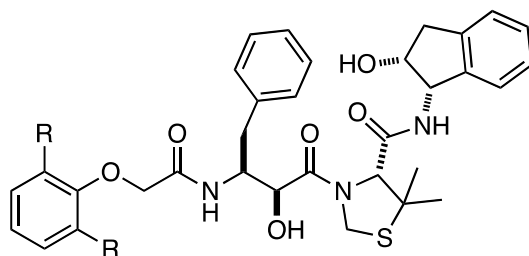
**Figure 1.10.4.** Water binding networks.<sup>cii</sup>



In a structural analysis of ligands bound to HIV-protease, Friere suggested that a cavity located in the binding pocket was not optimally filled by **1.98**.<sup>ciii,civ</sup> In attempts to explore the effects of filling this cavity with apolar substituents, they found that the

substitution of R = H (**1.98**) with R = F (**1.99**), Cl (**1.100**), and Me (**1.101**) all led to increases in binding affinities relative to **1.98** that were driven by increases in favorable binding enthalpy, not entropy. It was argued that the boost in *enthalpy* could be attributed to an increase in dispersive interactions arising from the larger vdW radii of the phenyl substituents, which also correlated with a more ordered structure in the binding pocket. The investigators suggested that the favorable entropies of binding arose from additional solvent desolvation that does correlate well with the increase in buried surface area.

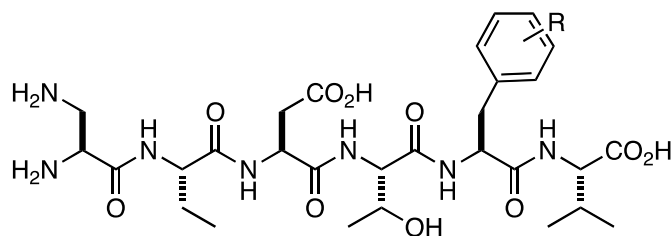
**Figure 1.10.5.** Freire ligands where binding cavity is filled via methyl substitution.



<i>Ligand</i>	R	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G^\circ$ (kcal/mol)	$\Delta H^\circ$ (kcal/mol)	$-T\Delta S^\circ$ (kcal/mol)
<b>1.98</b>	<b>H</b>	$8.5 \pm 0.7$	$-10.8 \pm 0.01$	$-4.5 \pm 0.2$	$-6.3 \pm 0.01$
<b>1.99</b>	<b>F</b>	$28.9 \pm 0.7$	$-11.5 \pm 0.1$	$-6.1 \pm 0.5$	$-5.4 \pm 0.1$
<b>1.100</b>	<b>Cl</b>	$192.3 \pm 0.2$	$-12.7 \pm 0.3$	$-6.4 \pm 0.4$	$-6.2 \pm 0.3$
<b>1.101</b>	<b>Me</b>	$714 \pm 0.4$	$-13.5 \pm 0.2$	$-6.5 \pm 0.7$	$-7.0 \pm 0.2$

Spaller and coworkers have also studied the effects of introducing halogen substitution on an aromatic ring of a PDZ domain agonists in a manner that increases the ligand hydrophobic surface area.<sup>cv</sup> They found that halogen substitution on phenyl rings increased binding affinity through an *entropically* favorable process (Table 1.10.5). This effect was attributed to the increase in the vdW radii of the halogens relative to an H atom at the *para* position of a phenyl ring (H: 1.20 Å, F: 1.47 Å, Cl: 1.77 Å).<sup>cvi</sup> Presumably, this would lead to the greater release of ordered water molecules that would give rise to favorable entropies. They observed that binding affinity increased upon the substituent heteroatoms with larger van der Waal radii in a process that was driven by a *favorable entropy*; however, they did not obtain any structural information of complexes in the binding pocket to confirm/support this hypothesis. Thus, although this is an interesting study on the binding thermodynamics of substituent effects on phenyl rings, the requisite structural data needed to determine the ligand buried surface area and possible dispersive interactions made in the binding complex were not obtained.

**Table 1.10.6** Spaller studies on substituent effects.



<i>Ligand</i>	<i>R</i>	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal mol <sup>-1</sup> )	$T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>1.102</b>	<b>4-H</b>	-8.0 ± 0.1	-6.9 ± 0.2	1.1 ± 0.2
<b>1.103</b>	<b>4-F</b>	-8.1 ± 0.1	-2.9 ± 0.5	5.2 ± 0.4
<b>1.104</b>	<b>4-Cl</b>	-8.6 ± 0.1	-4.3 ± 0.1	4.3 ± 0.1
<b>1.105</b>	<b>4-Br</b>	-8.5 ± 0.1	-4.4 ± 0.1	4.1 ± 0.2
<b>1.106</b>	<b>4-I</b>	-8.8 ± 0.1	-4.4 ± 0.1	4.4 ± 0.1

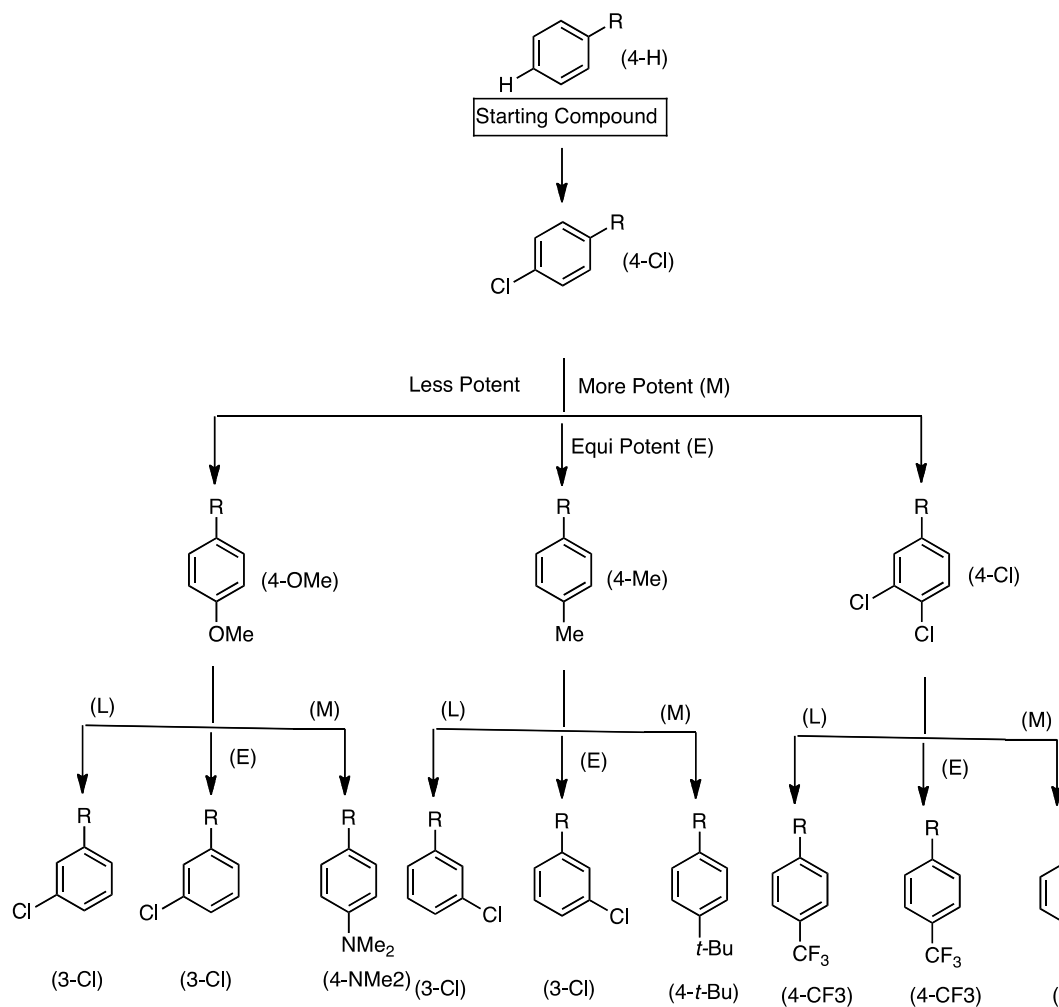
Similarly, the Whitesides group has recently investigated the effects of fluoro substituted methylene chains in potent binders of human carbonic anhydrase II (HCA II).<sup>cviii</sup> They observed that fluorinated derivatives bound with more favorable enthalpies and entropies relative to the parent alkyl chains. The more favorable enthalpy that was observed for the fluorinated derivatives has been proposed by the Whitesides group to arise from expulsion of added water in the binding pocket due to the larger van der Waal radii of the C-F bonds. It is noted that since the dispersive interactions would be weaker in C-F bonds than they are in C-H bonds, then there should be no added dispersive interactions arising from the fluorine mutation. It is speculated

that the larger radii of the fluoroalkyl group may displace more unoptimally hydrated water molecules that would then be hydrated in bulk water. However, their structural data did not provide insight regarding differences in hydration of complexes in the binding pocket to confirm/support this hypothesis, as accounting for water molecules in the binding complex using X-ray crystallography can be unreliable.

### **1.11 Topliss Operational Schemes.**

One of the earlier tactics that was applied to modify potential drug leads screening of libraries is the operational schemes of the Topliss decision tree. In this process, a whereby a small set of phenyl-substituted analogs are synthesized and their measured binding affinities guide the synthesis to a compound with optimized potency, as shown in Figure (1.11.1).<sup>cviii</sup> These choices in synthesis are guided by considerations of hydrophobicity ( $\pi$  values),<sup>cix</sup> steric effects,<sup>cx</sup> and electronic factors ( $\sigma$  values)<sup>cx</sup> of different substituents.

**Figure 1.11.1. Topliss operational scheme<sup>cx i</sup>**

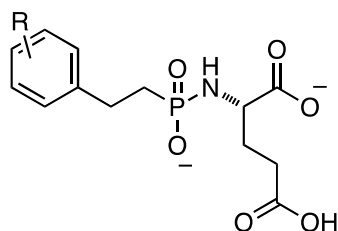


In order to apply this strategy, the first step involves the synthesis and testing of the initial lead. The selection of the next compound depends up whether or not the analog binds with more, equal or less potency. This process is repeated according to the Topliss decision tree (Figure 1.11.1). Ideally, one would need to synthesize only two to seven compounds in the early stages of drug design using the Topliss operational schemes to achieve maximal potency in the series. However, given today's resources and brute force efforts, many investigators will approach Topliss optimization by synthesizing the majority of the substituents shown in the schemes.<sup>cx ii</sup>

Interestingly, a literature search revealed that there has been no reported systematic study on the Topliss decision tree that reveal the binding thermodynamics entropy and enthalpy. There have also been very few studies that correlate changes in binding affinity with lipophilicity or the electronic properties of the respective substituents. Even more surprising is that there is no thorough study of the Topliss operational scheme that utilizes calorimetric methods to obtain the enthalpy and entropy of binding. However, in a study on increasing the potency of inhibitors for the prostate-specific membrane antigen (PSMA), Berkman et al. determined that ligand potencies correlated with the  $\pi$  and  $\sigma$  values of the suggested substituents of the Topliss scheme.<sup>cxiii</sup> As shown in Figure 1.11.2, the  $\pi$  and  $\sigma$  (each shown relative to R = H) are included to depict that the highest inhibition was met when both factors were sequentially increased. This led the authors to design a second set of inhibitors reflecting the observed trend, which guided them to synthesize and test the 4-Cl, 3-CF<sub>3</sub>; 4-CF<sub>3</sub>; 2,4-Cl<sub>2</sub>. Unfortunately the substituents led to no significant increases in PSMA inhibition. This result may be an indication that there are more complexities to the schemes than solely relying on  $\pi$  and  $\sigma$  values to target gains in binding affinity.



**Figure 1.11.2.** Berkman ligands.



Inhibitor	$\pi$ (R only)	$\sigma$ (R only)	% Inhibition
<b>1.102:</b> R = H	0	0	71
<b>1.103:</b> R = 4-Me	0.56	-0.17	74
<b>1.104:</b> R = 4-OMe	-0.02	-0.27	0
<b>1.105:</b> R = 4-Cl	0.71	0.23	85
<b>1.106:</b> R = 3,4-Cl <sub>2</sub>	1.25	0.52	100

## 1.12 Conclusions

Recent advances in calorimetry have allowed for the facile understanding of biomolecular interactions between small molecules and their host receptors to provide the entropy and enthalpy changes that occur upon a binding event with unprecedented accuracy. This advance in technology has brought with it a renewed interest in studying many design strategies used in the pharmaceutical industry to develop potent and selective inhibitors for medically relevant therapeutics.

One tactic to optimize affinity is to preorganize a flexible ligand into its bound conformation. It is assumed that rigidifying the ligand by reducing the total number of rotors will lower its entropy, thereby reducing the entropic penalty associated with binding. Although the entropy and enthalpy of binding in many case studies have not been determined, there have already been several cases where preorganization of a ligand does not grant the anticipated gains in binding affinity driven by a more favorable entropy. There is also little scientific evidence of a correlation between the reduction of ligand rotatable bonds and binding affinity in biomolecular interactions.

It is also believed that increasing the nonpolar surface area of a ligand may lead to enhanced binding affinity due to a favorable change in entropy. Such boosts in entropy are posited to arise from the release of highly ordered water molecules during ligand complexation. Although it has been shown increases in drug potency can be achieved by the addition of nonpolar surface area to a ligand, recent thermodynamic and structural studies have been reported suggesting that such affinity increase may be the product of favorable changes in *enthalpy* and not *entropy* driven. Since it is generally thought that the driving force hydrophobic effects were favorable changes in entropy, which as we have shown is not always observed, we wondered how these affinity enhancements are achieved and if they could be better understood and implemented.

We also reviewed common strategies for the optimization of drug leads containing phenyl substitution. A common motif that is used for optimizing the affinity of such ligands is the use of the Topliss decision tree. Although there have been hundreds of published studies on drug lead development employing this scheme, there is not a single study correlating its thermodynamic effects associated with its implementation in ligand optimization. Since such thermodynamic information is lacking in the current application of the schemes, we believe that it is imperative to investigate the Topliss decision tree to determine if any possible trends can be identified that may lead to a better understanding, and possibly implementation, of this optimization process.

In summary, we reviewed the current literature regarding common strategies and motifs employed by big pharma to discover potent drug leads. We dove into this task with a keen interest on questioning if the strategies that we encountered were explainable, predictable, and most importantly, do they work? Then, we critically examined

representative examples of significant strategies in drug design such as (1) ligand preorganization, (2) affinity enhancements via the increase of nonpolar surface area, and (3) the use of the Topliss decision tree. We found with overwhelming conviction that each of these design strategies were poorly understood because each is backed by *assumptions* that have little to no support from thermodynamic or structural evidence. Thus, we gathered the questions and observations from our literature review to plan a study wherein these design motifs are explored to better understand how and why they are effective, so that we could potentially improve how they are employed in the early stages of drug discovery. Specifically, does the restriction of ligand rotors in protein-ligand interactions corresponding to a favorable change in entropy of 1.1–1.5 kcal•mol<sup>-1</sup>? What is the thermodynamic driving force, if any, that increases affinity when apolar surface area is added to a drug lead. And, are there any thermodynamic trends of the Topliss decision tree that correlate with favorable gains in ligand potency?

## Chapter 2 Thermodynamic and Structural Evaluation Preorganizational Constraints for MUP-1 Ligands.

Earlier studies on ligand preorganization with several Grb2 SH2 domain-binding peptides revealed that cyclopropane-derived mimetics consistently bound with greater potency than their respective flexible controls due to *a more favorable enthalpic advantage, not an entropic one.*<sup>cxiv</sup> Since this was the first example wherein a preorganizational constraint led to an *enthalpically driven* increase in binding affinity relative to a flexible control, additional studies such as structural characterization<sup>cxv</sup> and computer-aided simulation<sup>cxvi</sup> of the binding event were obtained to discern the origin of this finding. However, we are still working towards understanding the energetics associated with the interactions of binding between peptide ligands and the Grb2 SH2 domain. Whilst pursuing those studies, we also sought another system where we could study the effects of ligand preorganization and hydrophobicity. Accordingly, we embarked on a study to further examine this phenomenon by the design, synthesis and biological evaluation of ligands that bind to the mouse Major Urinary Protein (MUP-I).

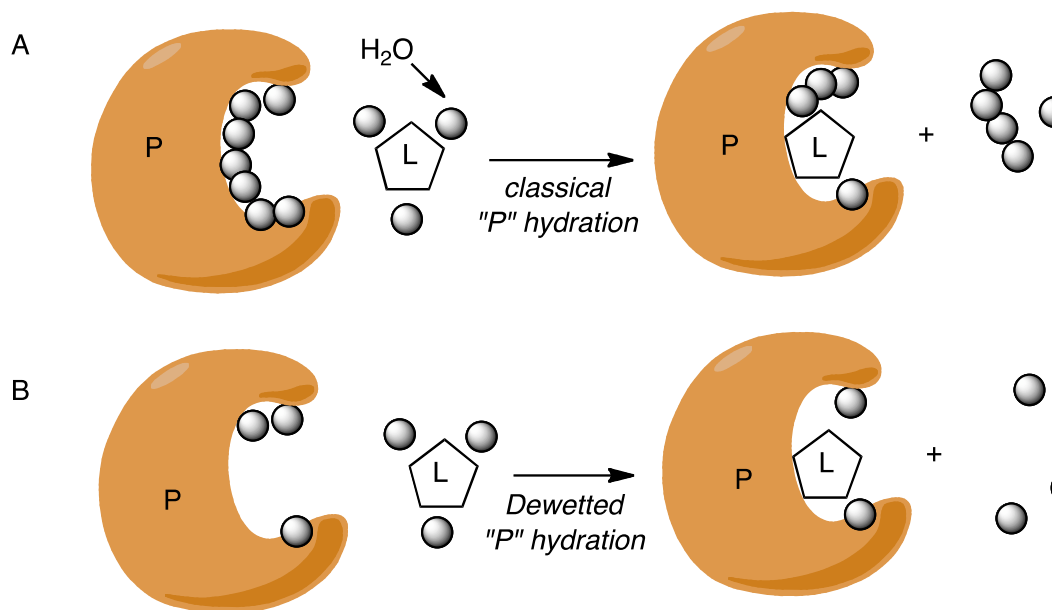
## 2.1 Mouse Major Urinary Protein I (MUP-I).

**Figure 2.1.1** Structure of MUP-I displaying the  $\beta$  barrel binding cavity.



The lipocalins are a small family of proteins that adopt an eight-stranded  $\beta$ -barrel cavity,<sup>cxvii</sup> and MUP-I is a member of this family and a pheromone transport protein in male rodent urine.<sup>cxviii,cxix</sup> It forms complexes with pheromones in the male mouse that are then excreted in the urine. The active binding site in MUP is enclosed so that there are no solvent exposed regions in the pocket (Figure 2.1.1). Interestingly, recent calculations have indicated that MUP-I has an unoptimally hydrated binding pocket that is dewetted.<sup>cxx</sup> These studies indicate that unlike a model protein system where water exists in the *apo* structure and is then expelled upon protein ligand binding, MUP-I has regions of the binding pocket that are not optimally hydrated by solvent. Thus as shown in Figure 2.1.2, it is speculated that solute-solute interactions in these dry regions may not be accompanied with expulsion of water upon binding to any significant extent.

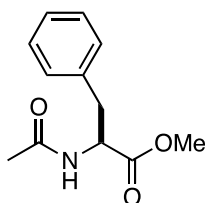
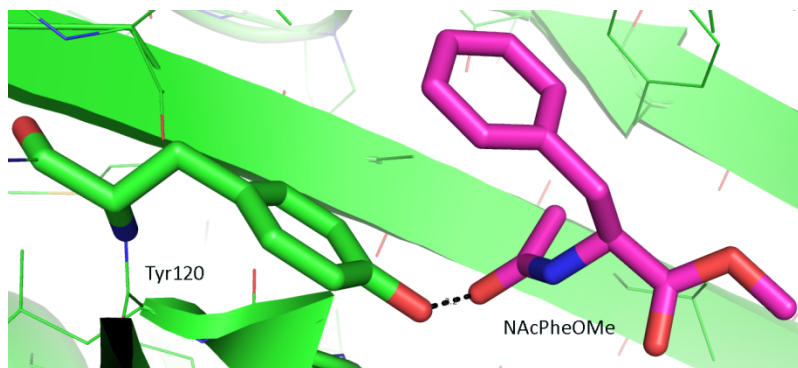
**Figure 2.1.2** Protein-ligand association a ligand and a “dewetted” protein binding domain. (A) generally accepted hydration of a protein and (B) the “dewetted” hydration state.



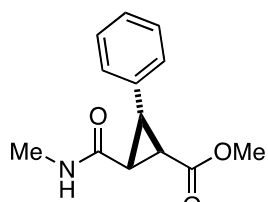
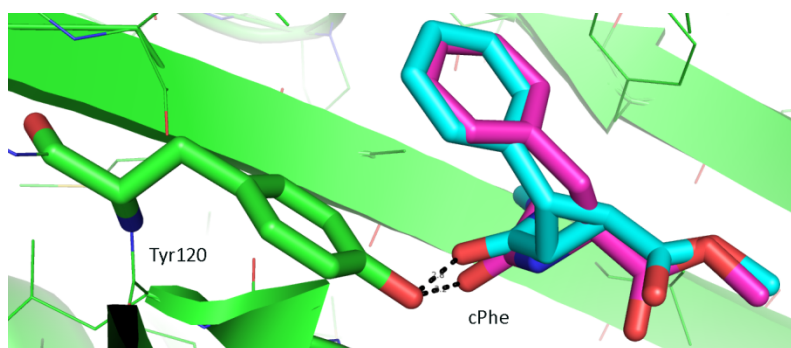
## 2.2 Ligand Design

It was disclosed several years ago in a personal communication from the Homans group that Phe and Ile derivatives were shown to be modest binders to MUP-I. X-ray structural analysis of these compounds were examined, and we queried whether a cyclopropane constraint similar to those used in ligands that bound to the Src and the Grb2 SH2 domains would preorganize the Phe derivative into its bound state (Figure 2.2.1). Given that the cyclopropane moiety in **2.08** adds a heavy atom to the ligand skeleton relative to **2.06**, we designed a suitable flexible control **2.07** that lacked the cyclopropane constraint (Figure 2.2.2).

**Figure 2.2.1** MUP-I ligand design featuring (a) the crystal structure of NHAcPheOMe in the protein bound complex and (b) and overlay of this structure with the proposed constrained mimic.



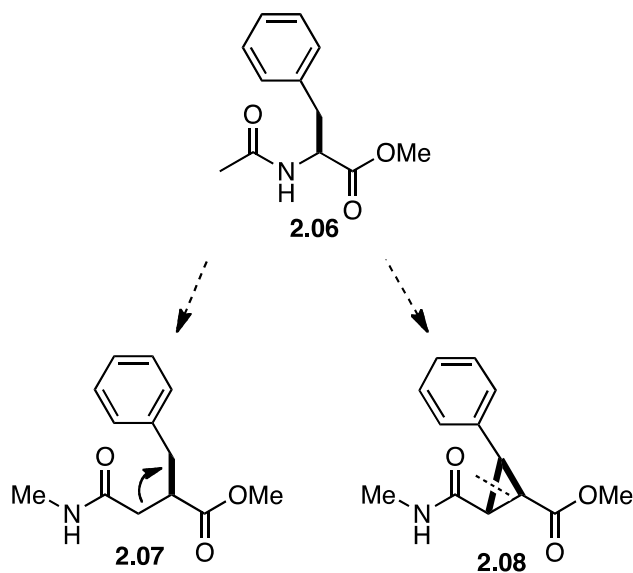
(a). **2.06**: AcNPheOMe



(b) **2.08**: AcNcPhe-OMe



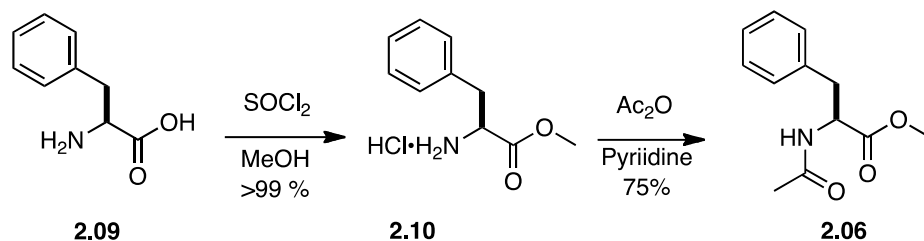
**Figure 2.2.2** Ligand design for studies with MUP-I.



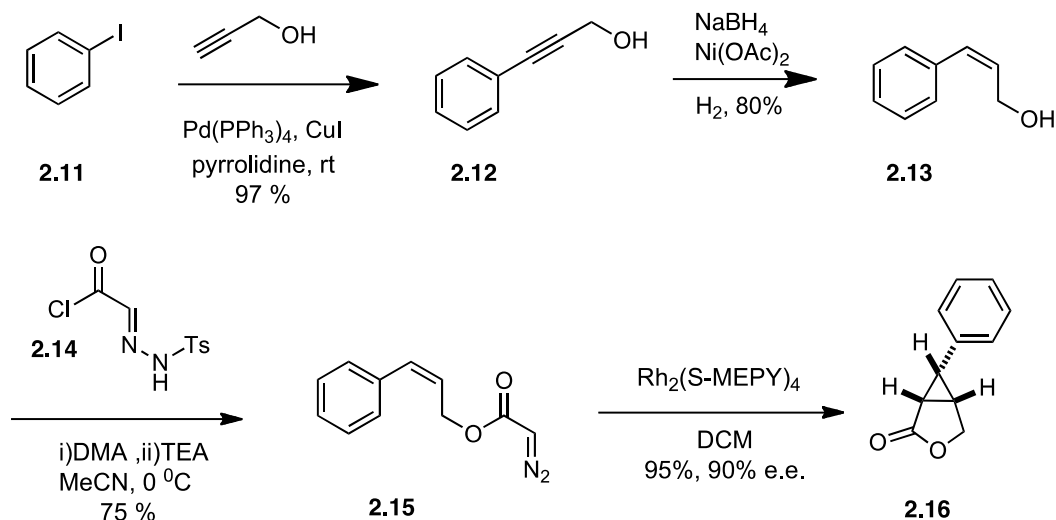
### 2.3 Ligand Synthesis

The ligand *N*-acetyl phenylalanine methyl ester, shown as AcNPheOMe in Figure 2.4.2, was prepared by the esterification of phenylalanine using thionyl chloride in methanol to provide an amine hydrochloride that was then acylated using pyridine and acetic anhydride.<sup>cxxi</sup> The cyclopropane derived Phe mimic **2.08** was synthesized using a known procedure for the synthesis of cyclopropyl lactones.<sup>cxxii,cxxiii</sup> The synthesis starts with a Sonogashira coupling of iodobenzene and propargyl alcohol to give **2.13** (Scheme 2.3.2), which was then selectively reduced in the presence of P-2 Ni. Treatment with the Corey-Meyers acid chloride<sup>cxxiv</sup> provided the diazoester **2.15** in good yield. For the key transformation of this synthesis, the diazoester was treated with  $\text{Rh}_2(\text{S-MEPY})_4$  to provide the cyclopropyl lactone **2.16** in 95% yield and 90% e.e (determined by chiral HPLC).

**Scheme 2.3.1 Synthesis of AcNPheOMe (2.06).**



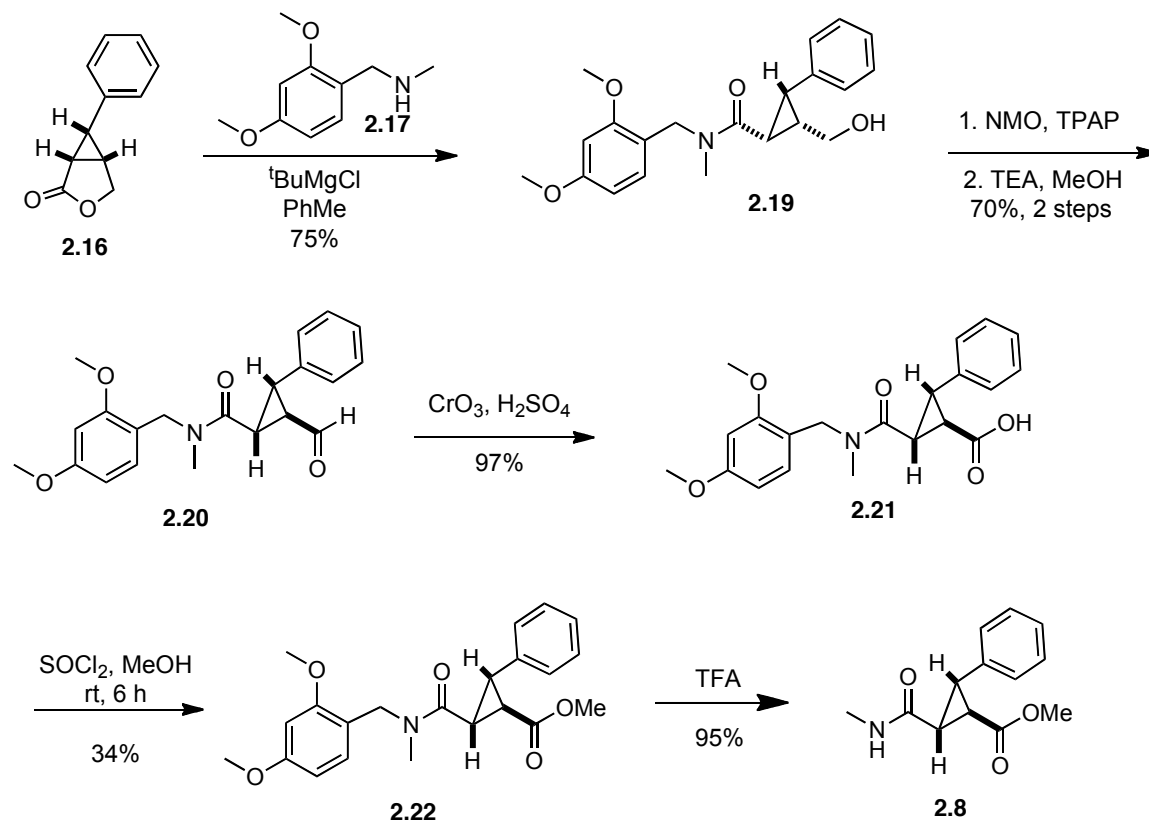
**Scheme 2.3.2 Synthesis of cyclopropane precursor.**



The known cyclopropyl lactone **2.16** was then opened with **2.17** to provide the DMB protected amide **2.19** (Scheme 2.3.3).<sup>cxxv</sup> Alternatively, **2.16** could be opened under mild conditions with **2.17** in the presence of  $\text{AlMe}_3$ <sup>cxxvi</sup>. This transformation was performed in hopes to increase reaction yield, but the procedure with  $\text{tBuMgBr}$  consistently gave higher yields. The alcohol **2.19** was then oxidized to an intermediate aldehyde that was epimerized to give **2.20**, which was then oxidized to the acid **2.21**. The

methyl ester of **2.21** was prepared by reaction with  $\text{SOCl}_2$  in MeOH, and the N-protecting group was removed using TFA to provide the final ligand **2.8**.

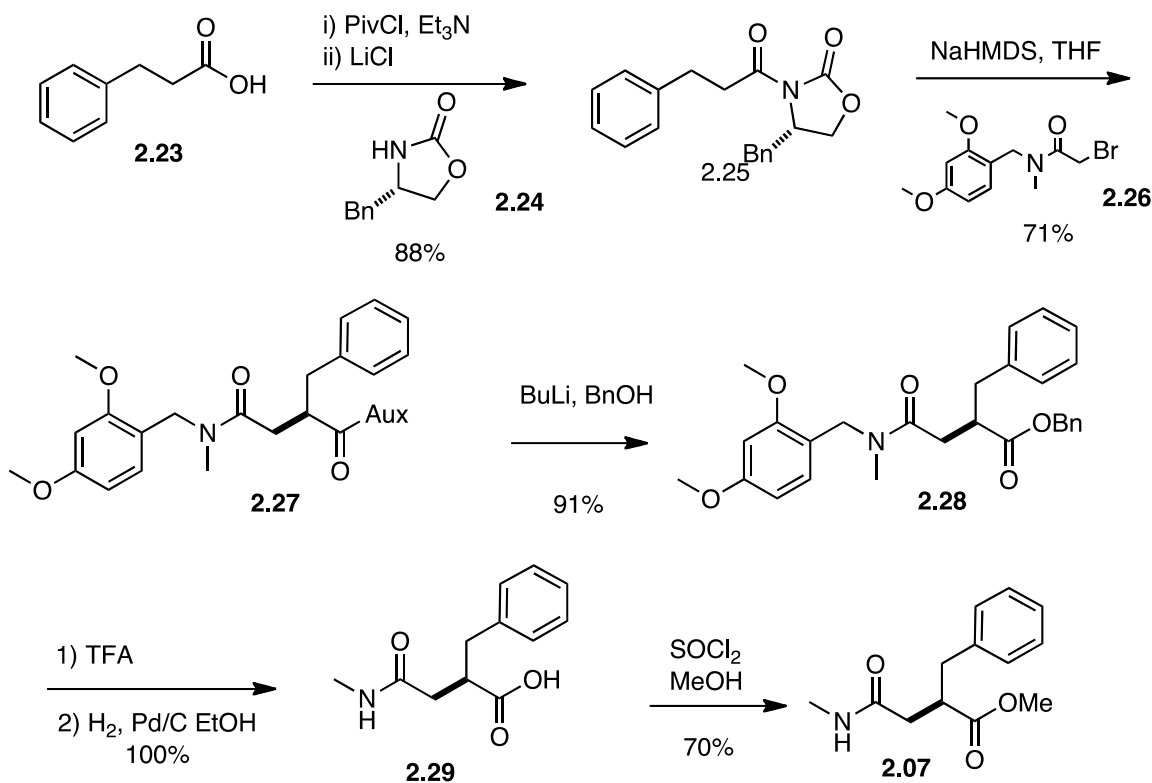
**Scheme 2.3.3** Synthesis of constrained mimic.



The flexible mimic **2.07** was prepared by colleague Dr. J. Tian as shown in Scheme 2.3.4. Briefly, **2.23** was treated sequentially with pivaloyl chloride and then the phenyl derived oxazolidinone **2.24** to give **2.25**, which was then treated with pivaloyl chloride to provide the intermediate mixed anhydride that was allowed to react with the phenyl derived oxazolidinone **2.24** to give **2.25**.<sup>cxxii</sup> Formation of the sodium enolate of **2.25**, followed by addition of **2.26** provided **2.27** as a single diastereomer in 71% yield. Removal of the chiral auxiliary from **2.27** using the lithium benzyloxide followed by

DMB deprotection using TFA afforded **2.29**, which was treated with thionyl chloride in methanol to give **2.07** in 70% yield.

**Scheme 2.3.4** Synthesis of flexible mimic



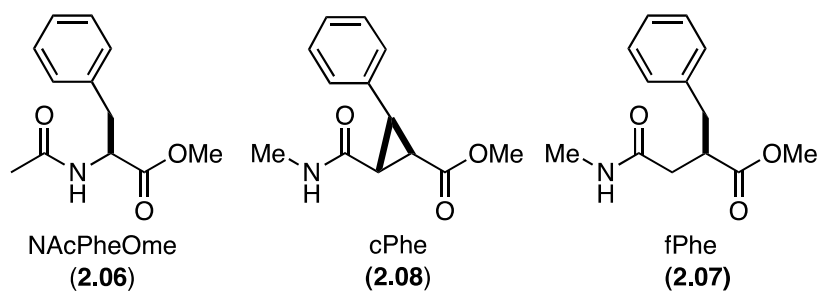
## 2.4 Thermodynamic Studies

With the constrained and flexible **2.08** and **2.07** in hand, we then conducted binding studies on these ligands with MUP-I (Table 2.4.1). It was found that the cyclopropane derived Phe mimic (cPhe) bound with nearly equal affinity to protected amino acid **2.06**. However, the constrained mimic **2.08** had a more favorable enthalpy of binding than the native ligand. Unfortunately, this gain in affinity was compensated by a relatively unfavorable entropy changes so that there was no significant change in binding

affinity. The flexible mimic **2.07** was then found to bind so weakly that the thermodynamics of binding could not be determined Table 2.4.1 because it had a  $\Delta H^\circ$  of binding that was too low for the instrument to detect.

It shall be noted that the suggested detection limit for the ITC is based off of the *enthalpy* of binding. Thus, it is possible that the flexible mimic could have a binding affinity that is driven by a strong *entropic* force, and such a curve would not be detected well by ITC. However, if we assume for this discussion that the binding affinity of **2.07** is  $< 1 \times 10^3 \text{ M}^{-1}$ , then this would indicate a preference for the flexible analog that is  $>200$  fold that of the flexible ligand. To date, this would be the largest affinity boost achieved using *cyclopropyl* constraints to the best of our knowledge. However, without any thermodynamic or structural characterization of the binding event with **2.07**, then we can not make any insightful conclusions that would shed light on preorganizational constraints.

**Table 2.4.1.** ITC binding studies Phe derivatives with MUP-I, at 298 K in PBS buffer at pH 7.4



	$K_a$	$\Delta G^\circ$	$\Delta H^\circ$	$-T\Delta S^\circ$
Ligand	( $M^{-1}$ )	( $kcal \cdot mol^{-1}$ )	( $kcal \cdot mol^{-1}$ )	( $kcal \cdot mol^{-1}$ )
2.06	$2.8 \times 10^5$	$-7.7 \pm 0.1$	$-11.3 \pm 0.7$	$3.6 \pm 0.1$
2.08	$2.4 \times 10^5$	$-7.2 \pm 0.1$	$-12.4 \pm 0.7$	$5.1.3 \pm 0.1$
2.07	Did Not Bind ( $K_a < 1 \times 10^3$ )			

## 2.5 Summary

In summary, we have successfully designed, synthesized, and tested a preorganized amino acid derivative containing a 1,2,3-*trisubstituted* cyclopropane that was designed to orient the amino acid into its bound conformation. This constrained mimic was determined by ITC to be equally potent to a Phe derivative. It was found that a more favorable enthalpy of binding was compensated by an equally unfavorable entropy compared to the native ligand. In order to properly ascertain the effects of the cyclopropane constraint, a flexible control containing the same number of heavy atoms was synthesized and tested. The binding enthalpy of the flexible compound was decreased by such a large extent that it could not be detected by ITC. Given the known

detection limits of the instrument, this indicates that the flexible ligand was at least 200 fold *less* potent than the constrained analog. Potency differences of this magnitude have never been observed for cyclopropane constraints in protein-ligand interactions. However without structural data of the flexible control, it is difficult to infer if the differences in ligand binding affinity arose from the ligand constraint or some other unknown complexity to binding.

## 2.6 The Binding Thermodynamics of Phenyl Alkanols

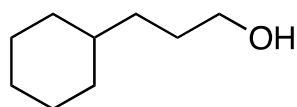
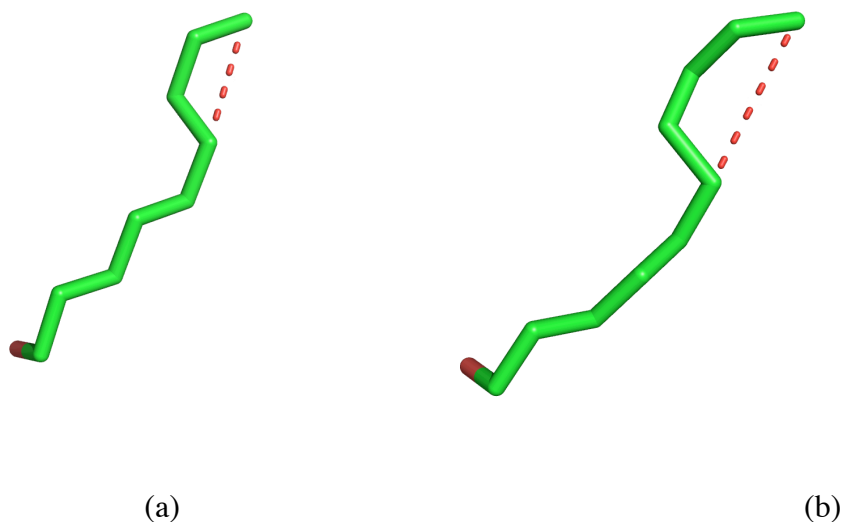
Thermodynamic studies from our investigation of trisubstituted *trans* cyclopropanes that bound to MUP-I showed an unprecedented preference for a constrained analog that had a  $K_a \sim 200$  fold better than the flexible amino acid mimic from which it was derived. However, we were left with little thermodynamic and structural information to glean from this discovery because no thermodynamic data could be obtained for the flexible analog. In order to further our investigation of the effects of ligand preorganization constraints on this system, we set out to modify the initial ligand design and to study novel binders that would be smaller, less substituted, and more readily available for derivative testing.

Our attention turned back to a class of linear alkanols that were shown to be potent MUP-I binders.<sup>cxxvii</sup> As shown in Section 1.10, the affinity of the *n*-alkanols for MUP-I increased in an enthalpically driven process upon the addition of each methylene group. Structural studies revealed that these alkanols adopted an extended conformation in the binding pocket until  $n = 4$  (nonanol) at which point the ligands adopted a turned

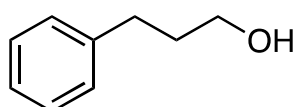
conformation (shown in Figure 2.6.1). This turned structure was maintained for the decanol derivative. We were intrigued by this finding and sought to study the effects of substituting in a cyclohexyl substituent to enforce its turned posed in the pocket. Preliminary studies on the cyclohexyl-substituted ligand **2.32** were somewhat disappointing because it was a weaker binder than we anticipated. However, we then replaced the cyclohexane moiety with a phenyl ring and found that the analog **2.33** had better binding based on preliminary runs.



**Figure 2.6.1.** Above: Binding pose of nonanol (a) and decanol (b) wherein it adopts a turned configuration (shown as a dash). Below: cyclohexyl (#) and phenyl (#) alcohols that were proposed to mimic the turn.



**2.32**



**2.33**

Based upon these preliminary observations, we opted to pursue a study of the binding of a series of phenyl alkanols by studying the effect of increasing nonpolar surface area by adding methylene groups. Because it had been shown that the alkanols adopted an extended, zig-zag conformation in the binding pocket (up until nonanol), we hoped to develop our phenyl alkanol system to study the effects of ligand preorganization using an internal olefin. Thus, we set out to study the binding thermodynamics of simple phenyl alkanol derivatives utilizing ITC to determine if this system would be beneficial for our aforementioned goals.

## 2.7 Developing the ITC protocol

Our initial attempts at binding studies of phenyl propanol plagued by the inability to obtain reproducible values for  $\Delta H^\circ$  and  $T\Delta S^\circ$ . The difficulties that we faced arose from a combination of unforeseen problems attributed to a low  $c$  value and inconsistent  $n$  values. It soon became evident that we needed to revamp our ITC protocols to obtain consistent and reproducible ITC data.

We first decided to correct the issues with our  $c$  value by doing a study wherein the concentration of protein and ligand were adjusted. This was carried out by using both the  $K_a$  from preliminary ITC results and a targeted  $c$  value of 500 which can then be inserted into eq (1.01) to solve  $[P]_t$ .

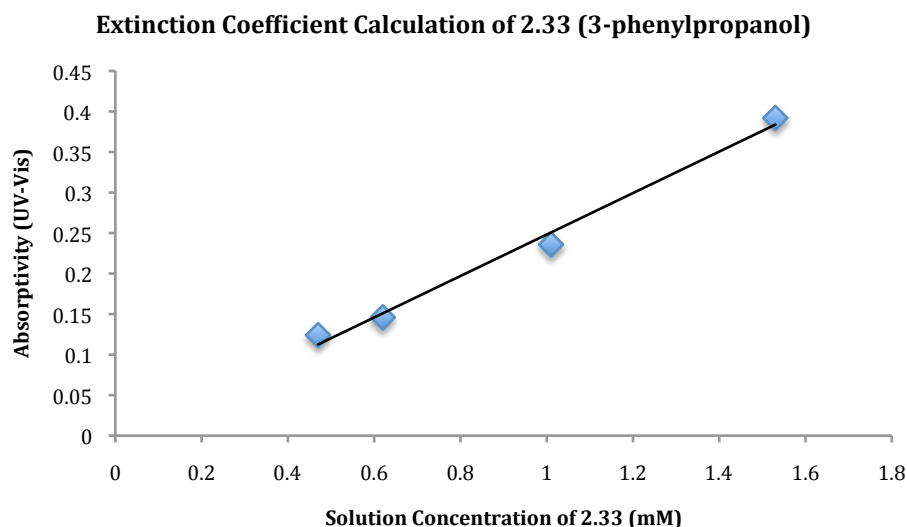
$$\frac{1}{r} = c = K_a[P]_t \quad (\text{eq 1.01})$$

The ligand concentration then had to be adjusted to obtain good signal to noise ratios for our thermograph, and this was done by maintaining a 12-fold equivalence of  $[L]$  to  $[P]$ . We thus identified ligand and protein concentrations that consistently produced isotherms having acceptable  $c$  values, but our  $n$ -values were still significantly different from 1.

It has been suggested that  $N$  values that do not reflect the supposed binding stoichiometry, or are irreproducible, can be attributed to errors in: (1) the protein concentration, (2) ligand concentration or (3) ligand occupancy in the complex during the binding event. We opted to address these issues by first assuming that the ratio of ligand to protein was 1:1, and this was considered a safe assumption at that time since a number of related alkyl alkanols were shown to bind with MUP-I in a 1:1 ratio. This hypothesis

was later confirmed by X-ray crystallographic data. Therefore, we anticipated  $n$  values that were near 1.0, yet we consistently observed values around 0.7, and they ranged from 0.5 to 1.4. As a part of our protocol, the protein concentrations were routinely checked by UV-Vis spectroscopy using a known extinction coefficient  $\epsilon_{280} = 14,661 \text{ M}^{-1}\text{cm}^{-1}$  for the protein,<sup>cxvii</sup> but the ligand concentration was being measured by weight. Since both the ligand and protein had a chromophore, we measured the extinction coefficient of **2.33** (3-phenylpropanol) in the ITC buffer at pH 7.4 as shown in Figure 2.7.1. We then used this extinction coefficient to determine our ligand concentrations and most importantly how the UV and weight concentrations compared to each other. If the UV concentration matched the weight concentration within 5% error, we opted to proceed with the run using the UV corrected ligand concentration. Gratifyingly, this led to achieving isotherms that reproducibly provided consistent values for enthalpy and entropy, but our  $n$  values were still consistently less than 1.0. Since we had confirmed our ligand concentrations using UV-Vis spectroscopy and at this time knew that the ratio of protein to ligand in the binding pocket was 1:1, we concluded that the low  $n$  values were a result of an erroneous protein concentration in solution.




**Figure 2.7.1** Extinction coefficient calculation



Since it was previously shown that alkyl alkanols of the form  $\text{CH}_3-(\text{CH}_2)_n\text{-OH}$ , where  $n = 4\text{-}8$ , bound to MUP-I in an enthalpically driven binding event, we first determined the effects of adding methylene groups to *phenyl* alcohols of the form  $\text{Ph}-(\text{CH}_2)_n\text{-OH}$  where  $n = 3\text{-}5$ . These ligands (**2.33**, **2.34** and **2.35**) are commercially available and were distilled prior to their use in ITC, the thermodynamic parameters of binding for these alcohols were collected and are shown in Figure 2.7.2, where the comparison between UV and weight for each respective run is also shown. We were gratified to observe that the newly discovered series of phenyl alcohols bound to MUP-I with the same trend as the linear alcohols. That is to say, the addition of each methylene group led to a favorable increase in binding affinity that was driven by  $\Delta H^\circ$  not  $\Delta S^\circ$ . Variations in the enthalpy gains and entropy losses compared to those observed for the alkyl alkanols were observed and will be discussed in greater detail in the following section. The comparison of UV/weight

ligand concentrations for the data of our modified ITC protocol shows that variations in our  $n$  values did not result in variations of the thermodynamic parameters  $\Delta H^\circ$  and  $\Delta S^\circ$ , which suggests that variations in  $n$  values are due to protein concentrations and thus need to be corrected.

**Figure 2.7.2.** Individual run data analysis of MUP-I ligands at 298 K in PBS buffer at pH 7.4.

	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-\Delta S$ (kcal•mol <sup>-1</sup> )	N (UV)	N (wgt)
 <b>2.33</b>	-6.9	-12.2	5.3	0.71	0.71
	-6.9	-12.4	5.5	0.68	0.76
	-6.9	-12.3	5.5	0.57	0.60
<hr/>					
 <b>2.33</b>	-7.5	-13.6	6.1	0.66	0.64
	-7.4	-14.2	6.8	0.56	0.56
<hr/>					
 <b>2.33</b>	-8.3	-16.5	8.2	0.64	0.70
	-8.4	-16.2	7.7	0.71	0.71
	-8.2	-16.9	8.7	0.67	0.75

## 2.8 The Utility of the Internal Olefin to Restrict Rotors

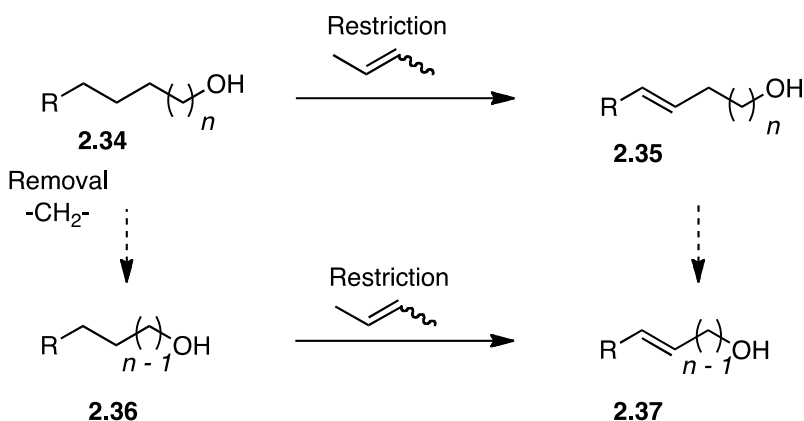
MUP-I is known as a “promiscuous binder” because it can associate with many different types of molecules, albeit they are typically small, hydrophobic ligands. Furthermore, typical binding processes are characterized by thermodynamics with a strong enthalpic driving force that are postulated to arise from dispersive interactions made in the binding complex.<sup>cxxviii</sup> Given our recently developed series of novel MUP-I binders of the form Ph-(CH<sub>2</sub>)<sub>*n*</sub>-OH where  $n = 3-5$ , we decided to build upon what was

known about MUP-I protein ligand interactions by revisiting our initial approach of evaluating the structural and thermodynamic effects of restricting rotors.

As mentioned in the Section 1.5, restriction of ligand rotors are estimated to provide reclaimable entropies between 1.1 – 1.5 kcal•mol<sup>-1</sup> per independent rotor restricted, based on theory and some notable observations.<sup>cxxviii</sup> Despite a lack of consistent experimental data, the predicted penalties for “freezing” a rotor have an impact upon key decisions in early drug development. Many scoring functions apply these values as context-independent additive per-rotor penalties<sup>cxxix</sup> such that these programs disfavor ligands with more rotors by simply counting the number of rotatable bonds. This practice of applying unbiased per-rotor penalties in computer-assisted ligand screening has recently been challenged on the basis that there is no evidence showing how rotatable bonds may correlate to losses of configurational entropy upon complex formation.<sup>cxxx</sup>

We herein assess the use of these so-called “context-independent” per-rotor penalties utilizing a systematic experimental approach. As shown in Figure 2.8.1, we wanted to examine the thermodynamic and structural effects of restricting a rotor utilizing an internal olefin on simple alkanols and phenylalkanols. We further extended this study by examining methylene group deletion on the resultant alka(e)nols. Although “deletion” of a CH<sub>2</sub> also constitutes as the removal of a heavy atom, the two transformations are seen as “equivalent” for many *in silico* screening programs,<sup>cxxix</sup> thus we felt that this system was well-suited to study this supposed equivalency upon removing rotors.

**Figure 2.8.1** Proposed restrictor study.



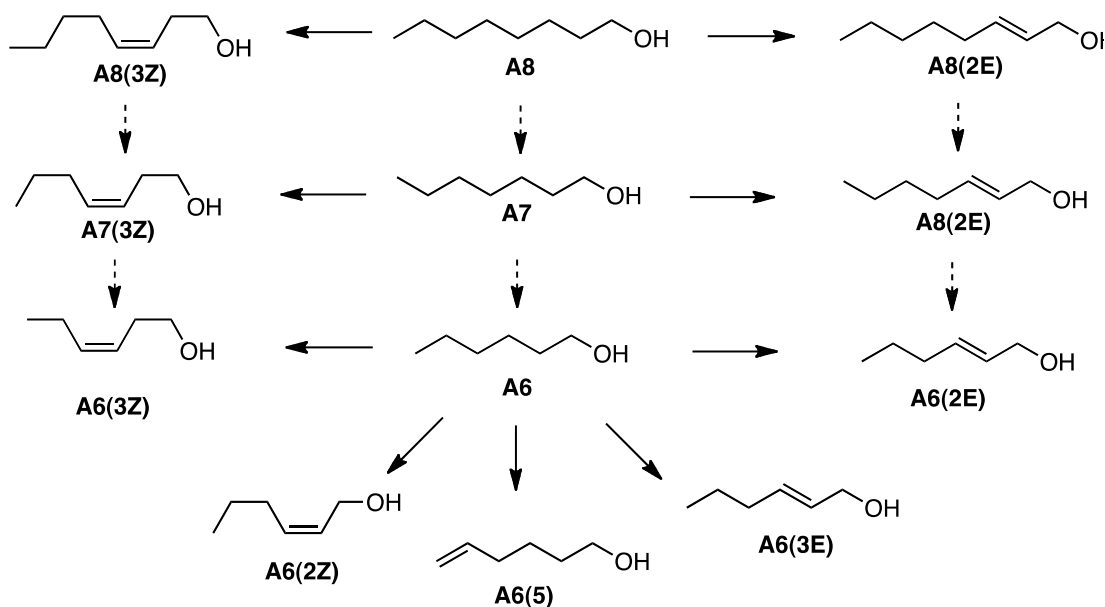
The proposed study therefore allows us to explore/ask to significant questions: Firstly, does reducing the number of rotors result in a change to ligand binding entropy by an average of 1.1 – 1.5 kcal•mol<sup>-1</sup>? Secondly, is there a binding entropy equivalence between reducing rotor numbers in an alkyl chain through restriction via introducing double bonds or removal through removing methylene groups? To explore this question, we collaborated with colleagues from the Homans lab at Leeds University. This granted us the ability of examining the variability of the R substituent shown in Figure 2.8.1 to include both phenyl and alkyl substituents, which gave us a diverse and larger number of substrates. We then synthesized and tested ligands containing an internal C-C double bond (**2.35**), and then removed a methylene group from this scaffold to obtain **2.37**, which will also be tested. Furthermore, since many of the alkyl alkanols were previously shown to bind in a zig-zag conformation, we also opted to test the Z olefin to probe the effect of olefin geometry on the entropy of binding.

Thus, the alkyl and phenyl alkanols shown in Figures 2.8.3 a and b were targeted for synthesis and biological testing. The Homans group conducted the ITC and X-ray of all

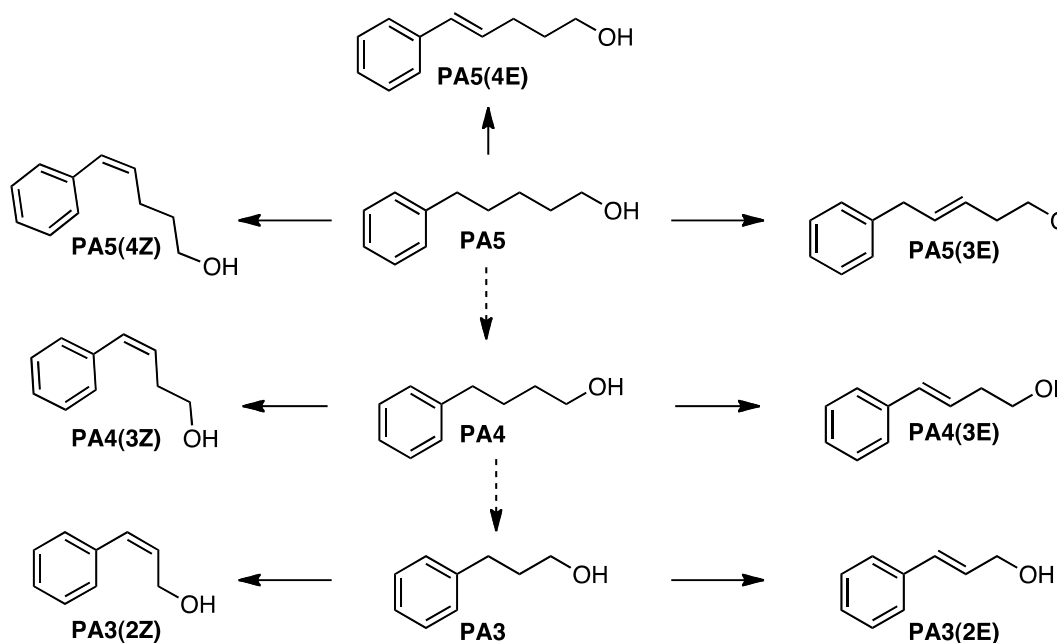
alka(e)nols, whereas we performed the synthesis, ITC, and X-ray analyses of all phenyl alka(e)nols.



**Figure 2.8.3.** Saturated and unsaturated *n*-alka(e)nols (a) and phenylalka(e)nols (b). <sup>a</sup>



**(a)** <sup>a</sup>Solid arrows depict introduction of a double bond while dashed arrows depict removal of a methylene unit. Abbreviations: “A” = alkanol, “PA” = phenylalkanol. “2Z” = Z-olefin at 2-position. “3E” = E-olefin at 3-position.

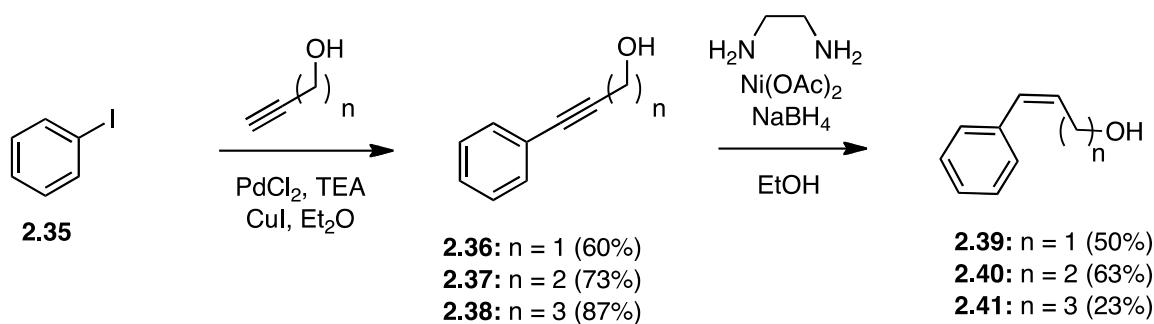


**(b)**

## 2.9 Ligand synthesis

Ligands **P6**, **P5**, **P4**, **P3** and **PA3(2E)** were from commercial sources. The syntheses of Z-olefins **PA3(2Z)**,<sup>cxxxix</sup> **PA4(3Z)**,<sup>cxxxix</sup> **PA5(4Z)**,<sup>cxxxix</sup> were initiated by coupling iodobenzene (**2.35**) with the requisite acetylides via a Sonogashira coupling to afford **2.36-2.38** (Scheme 2.9.1). These acetylides were selectively reduced utilizing P-2 Nickel, which was generated *in situ* from Ni(OAc)<sub>2</sub> and NaBH<sub>4</sub> to afford the requisite Z-olefins **2.39-2.41** in moderate to good yields. It should be noted that Lindlar conditions were also examined for this transformation, but it was often times too slow and despite extended reaction time, running without the presence of quinoline, and utilizing a Parr Shaker, all attempts at this transformation led to mostly unreacted starting materials.

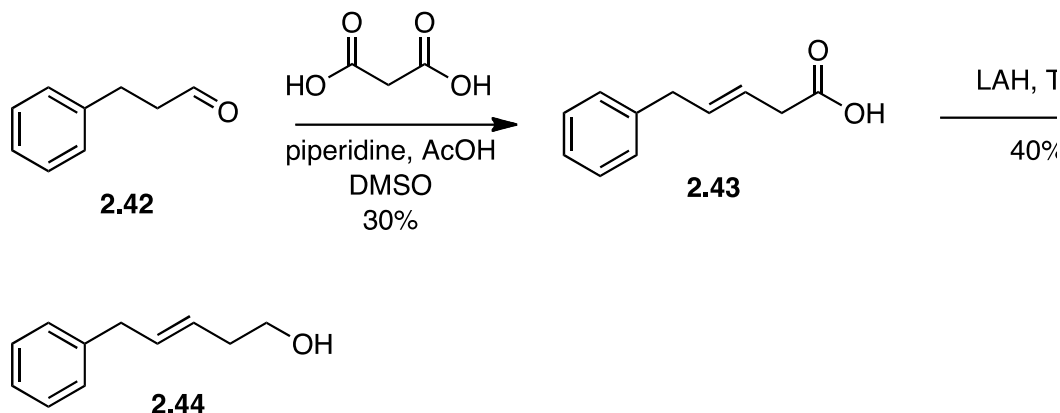
**Scheme 2.9.1** Synthesis of Z olefin.



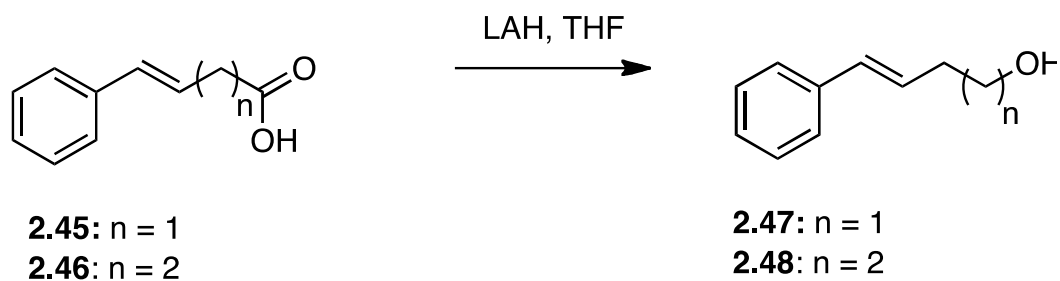
Phenyl alkenol **PA5(3E)** was synthesized starting with a Knoevenagel reaction wherein hydrocinnamyl aldehyde was treated with malonic acid in the presence of piperidine and AcOH to afford the **2.43** acid in 30% yield (Scheme 2.9.2).<sup>cxxxix</sup> Treatment of this acid with LAH afforded the alcohol **2.44** in 40% yield. The remaining *E*-olefins

were made through reduction of their commercially available acid precursors. Accordingly, **2.45** and **2.46** were reduced using LAH to provide the alcohols **PA4(3E)<sup>cxxxiv</sup> (2.47)** and **PA5(4E) (2.48)** in 65% and 47% yields, respectively (Scheme 2.9.3).

**Scheme 2.9.2.** Synthesis of phenyl alkenol **2.44**



**Scheme 2.9.3.** E-olefin synthesis



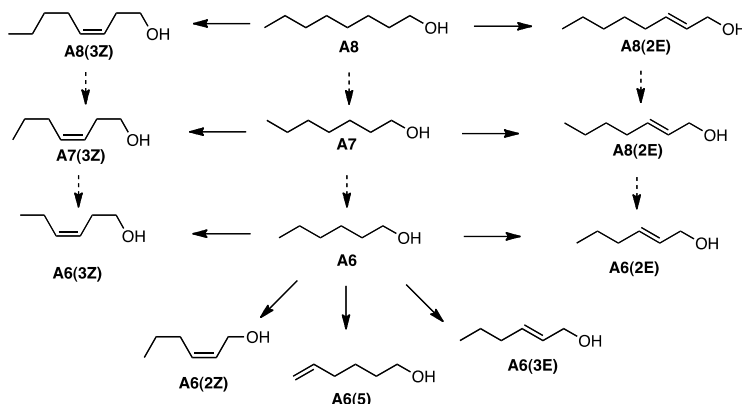
## 2.10 Thermodynamic Binding Studies of Alkyl and Phenyl Alka(e)nols

The observed thermodynamic parameters of binding,  $\Delta G_{\text{obs}}^{\circ}$ ,  $\Delta H_{\text{obs}}^{\circ}$ , and  $T\Delta S_{\text{obs}}^{\circ}$ , were measured by ITC at 298 K for the *n*-alka(e)nol and phenylalka(e)nols, and

these results are shown in tables 2.10.1 and 2.10.2, respectively. Dr. Richard Malham (R.M.) at Leeds University performed all ITC work for the alkyl alkanols and we performed all of the binding studies for the phenyl alka(e)nols. We obtained the ITC data for the phenyl alkanols using the established protocol where the protein concentration is normalized based on ligand concentration.

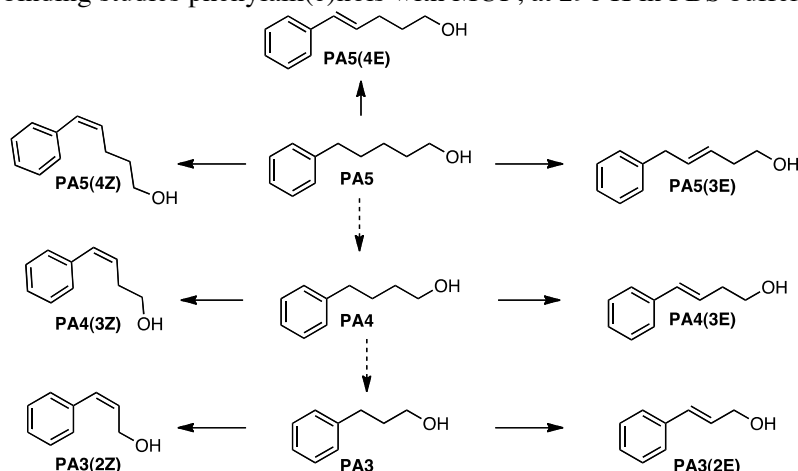
The thermodynamics of binding for ligands **A6**, **A7**, and **A8** were previously reported by the Homans group,<sup>cxvii</sup> and they are shown here *with* their respective alkenols which were previously unknown. Accordingly, we observed for both the alkyl and phenyl alka(e)nols that the binding affinity correlated well with the number of methylene groups in the alkyl chain. Therefore, ligands **A7**, **A72E** and **A73E** have a similar affinity for MUP-I that is different from the respective six and eight carbon analogues. The phenyl alkanols show the same trend, such that the ligands **PA3**, **PA3(2E)** and **PA3(2Z)** all have similar binding affinities within error.

**Table 2.10.1.** ITC binding studies alkylalkanols with MUP, at 298 K in PBS buffer, pH 7.4, performed by Richard Malham.



<i>Ligand</i>	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>A6</b>	$-6.2 \pm 0.07$	$-11.7 \pm 0.3$	$5.0 \pm 0.3$
<b>A6(2E)</b>	$-6.5 \pm 0.02$	$-10.6 \pm 0.4$	$4.1 \pm 0.4$
<b>A6(2Z)</b>	$-6.2 \pm 0.02$	$-9.7 \pm 0.8$	$3.5 \pm 0.8$
<b>A6(3E)</b>	$-6.3 \pm 0.02$	$-9.4 \pm 0.07$	$3.1 \pm 0.4$
<b>A6(3Z)</b>	$-6.2 \pm 0.02$	$-9.3 \pm 0.3$	$3.1 \pm 0.3$
<b>A6(5)</b>	$-6.1 \pm 0.02$	$-10.9 \pm 0.02$	$4.8 \pm 0.02$
<b>A7</b>	$-7.9 \pm 0.05$	$-13.4 \pm 0.4$	$5.5 \pm 0.4$
<b>A7(2E)</b>	$-8.2 \pm 0.5$	$-12.7 \pm 0.9$	$4.5 \pm 0.5$
<b>A7(3Z)</b>	$-7.2 \pm 0.02$	$-10.0 \pm 0.05$	$2.8 \pm 0.07$
<b>A8</b>	$-8.8 \pm 0.1$	$-14.5 \pm 0.5$	$5.6 \pm 0.2$
<b>A8(2E)</b>	$-8.5 \pm 0.05$	$-12.6 \pm 0.3$	$4.1 \pm 0.5$
<b>A8(3Z)</b>	$-8.1 \pm 0.02$	$-11.8 \pm 0.1$	$3.8 \pm 0.1$

**Table 2.10.2.** ITC binding studies phenylalk(e)nols with MUP, at 298 K in PBS buffer, pH 7.4<sup>a</sup>



<i>Ligand</i>	<i>n</i>	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>PA3</b>	0.71	$-6.9 \pm 0.02$	$-12.2 \pm 0.02$	$5.3 \pm 0.02$
<b>PA3(2E)</b>	0.50	$-6.8 \pm 0.02$	$-11.4 \pm 0.4$	$4.6 \pm 0.1$
<b>PA3(2Z)</b>	0.80	$-6.7 \pm 0.02$	$-8.4 \pm 0.5$	$1.7 \pm 0.1$
<b>PA4</b>	0.66	$-7.4 \pm 0.02$	$-13.8 \pm 0.7$	$6.4 \pm 0.07$
<b>PA4(3E)</b>	0.75	$-7.2 \pm 0.02$	$-12.0 \pm 0.02$	$4.8 \pm 0.02$
<b>PA4(3Z)</b>	0.75	$-7.3 \pm 0.02$	$-11.0 \pm 0.6$	$3.7 \pm 0.07$
<b>PA5</b>	0.71	$-8.3 \pm 0.02$	$-16.5 \pm 0.05$	$8.2 \pm 0.05$
<b>PA5(3E)</b>	0.85	$-7.9 \pm 0.5$	$-12.7 \pm 0.4$	$4.8 \pm 0.5$
<b>PA5(4E)</b>	0.84	$-8.3 \pm 0.02$	$-13.4 \pm 0.5$	$5.2 \pm 0.07$
<b>PA5(4Z)</b>	0.85	$-7.7 \pm 0.02$	$-13.1 \pm 0.7$	$5.4 \pm 0.02$

<sup>a</sup> *n*-values were normalized to upon adjusting the protein concentration to the known ligand concentration

*Observed Thermodynamics of Binding: Methylene Removal.* The difference in the observed thermodynamic parameters of binding,  $\Delta\Delta G^\circ_{\text{obs}}$ ,  $\Delta\Delta H^\circ_{\text{obs}}$ , and  $T\Delta\Delta S^\circ_{\text{obs}}$ , between each member of the *n*-alka(e)nol and phenylalka(e)nol series and each analogous

member that possesses one fewer methylene groups, which are the energetic consequences of methylene removal, are given in Tables 2.10.3 and 2.10.4, respectively. For the alkyl alkanols, we observed that a removal of a methylene group on average eventuates in a slightly favorable change in  $T\Delta\Delta S^{\circ}_{\text{obs}}$  of  $0.2 \text{ kcal}\cdot\text{mol}^{-1}$  that is dominated by an unfavorable  $\Delta\Delta H^{\circ}_{\text{obs}}$  of  $1.2 \text{ kcal}\cdot\text{mol}^{-1}$  such that  $\Delta\Delta G^{\circ}_{\text{obs}}$  for the alchemical modification is unfavorable by  $1.0 \text{ kcal}\cdot\text{mol}^{-1}$ . Similarly, we observed for the phenylalka(e)nols that the removal of a methylene group across the series eventuates in favorable binding entropies that are more than compensated by unfavorable binding enthalpies such that  $\Delta\Delta G^{\circ}_{\text{obs}}$  is  $0.6 \text{ kcal}\cdot\text{mol}^{-1}$ .

Next, we compared the two data sets of phenyl and alkylalka(e)nols. We found that on average the  $\Delta\Delta G^{\circ}_{\text{obs}}$  for the removal of a methylene group from the *n*-alka(e)nol series is more unfavorable by  $0.4 \text{ kcal}\cdot\text{mol}^{-1}$  than for the phenyl alka(e)nols because the average  $T\Delta\Delta S^{\circ}_{\text{obs}}$  of the former is less favorable by a statistically significant  $0.9 \text{ kcal}\cdot\text{mol}^{-1}$  than that for the latter. The average  $\Delta\Delta H^{\circ}_{\text{obs}}$  for removing a methylene from the *n*-alka(e)nols is less unfavorable than that for the phenylalka(e)nols, but these values are not significantly different as their error ranges overlap. That  $T\Delta\Delta S^{\circ}_{\text{obs}}$  for the removal of a methylene group from the *n*-alka(e)nols is significantly less favorable than for the phenyl alka(e)nols may reflect differences in the energetics of ligand desolvation, which will be further discussed in a following section. Lastly, we observed that the binding free energies of the ligand pairs **A6/PA3**, **A7/PA4**, and **A8/PA5** are similar, suggesting that a phenyl ring is the approximate equivalent of a trimethylene group with regard to binding affinity for this biological system.

**Table 2.10.3.** Difference in observed ITC for the removal of a methylene group in alkylalkanols.

<i>Modification</i>	$\Delta\Delta G_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )
<b>A8→A7</b>	<b>0.9 ± 0.1</b>	<b>1.0 ± 0.1</b>	<b>-0.1 ± 0.2</b>
<b>A7→A6</b>	<b>1.1 ± 0.4</b>	<b>1.6 ± 0.3</b>	<b>-0.5 ± 0.3</b>
<b>A8(2E)→A7(2E)</b>	<b>0.3 ± 0.5</b>	<b>-0.1 ± 4.1</b>	<b>0.4 ± 0.5</b>
<b>A7(2E)→A6(2E)</b>	<b>1.7 ± 0.5</b>	<b>2.1 ± 4.3</b>	<b>-0.4 ± 0.7</b>
<b>A8(3Z)→A7(3Z)</b>	<b>0.9 ± 0.05</b>	<b>1.8 ± 0.7</b>	<b>-0.9 ± 0.2</b>
<b>A7(3Z)→A6(3Z)</b>	<b>1.0 ± 0.05</b>	<b>0.7 ± 1.3</b>	<b>0.3 ± 0.3</b>
<b>Average</b>	<b>1.0 ± 0.5</b>	<b>1.2 ± 1.0</b>	<b>-0.2 ± 0.2</b>



**Table 2.10.4.** Difference in observed ITC for the removal of a methylene group in phenylalkanols.

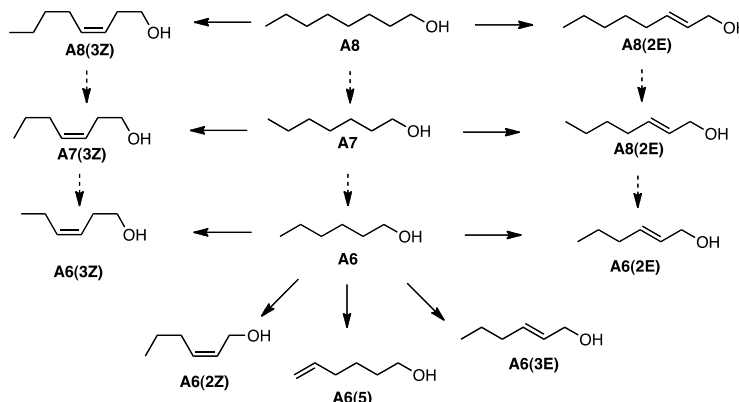
	$\Delta\Delta G_{obs}^0$	$\Delta\Delta H_{obs}^0$	$-\Delta T\Delta S_{obs}^0$
<i>Ligand</i>	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )
<b>PA5→PA4</b>	<b>0.9 ± 0.07</b>	<b>2.7 ± 1.0</b>	<b>-1.8 ± 1.3</b>
<b>PA4→PA3</b>	<b>0.6 ± 0.02</b>	<b>1.6 ± 0.8</b>	<b>-1.1 ± 0.4</b>
<b>PA5(4E)→PA4(3E)</b>	<b>1.0 ± 0.02</b>	<b>1.4 ± 0.9</b>	<b>-0.4 ± 0.4</b>
<b>PA4(3E)→PA3(2E)</b>	<b>0.4 ± 0.05</b>	<b>0.7 ± 0.8</b>	<b>-0.3 ± 0.1</b>
<b>PA5(4Z)→PA4(3Z)</b>	<b>0.4 ± 0.02</b>	<b>2.1 ± 0.9</b>	<b>-1.7 ± 0.07</b>
<b>PA4(3Z)→PA3(2Z)</b>	<b>0.7 ± 0.02</b>	<b>2.7 ± 3.2</b>	<b>-2.0 ± 0.1</b>
<b>Average</b>	<b>0.6 ± 0.02</b>	<b>1.8 ± 1.4</b>	<b>-1.2 ± 0.07</b>

In summary, we observed for the phenylalka(e)nols and alkylalka(e)nols that methylene group removal is accompanied by a dominating penalty in binding *enthalpy* that is slightly offset by an enhanced *entropy*. This is presumably due to the occurrence of a dewetted, hydrophobic binding pocket where nonpolar, nondirectional dispersive interactions dominate the binding process.<sup>cxxvii, cxxxv</sup> Similar trends in binding thermodynamics were reported by Stone and coworkers upon the addition of methylene groups to thiazole derivatives which bind to MUP-I.<sup>xcix</sup> They also observe that a favorable change in binding entropy is dominated by an unfavorable change in binding enthalpy such that the overall process is unfavorable for methylene group removal.

However, these results stand in contrast to several systems in which the removal/addition of methylene groups were studied. For example, Engberts et. al. previously studied the effect of incremental removal of methylene units to benzamidine chloride inhibitors on the thermodynamics of their binding to trypsin. They found that such removal of methylene groups to benzamidine chloride inhibitors to trypsin led to a decrease in affinity that was accompanied by an unfavorable entropy.<sup>cxxxvi</sup> It is also worth noting Whitesides' studies on carbonic anhydrase II binders that also varied in the number of methylene groups.<sup>cxxxvii</sup> Therein, it was determined that the removal of methylene groups to benzenesulfonamide derivatives resulted in losses in both binding entropy and enthalpy such that the overall binding affinity decreased.

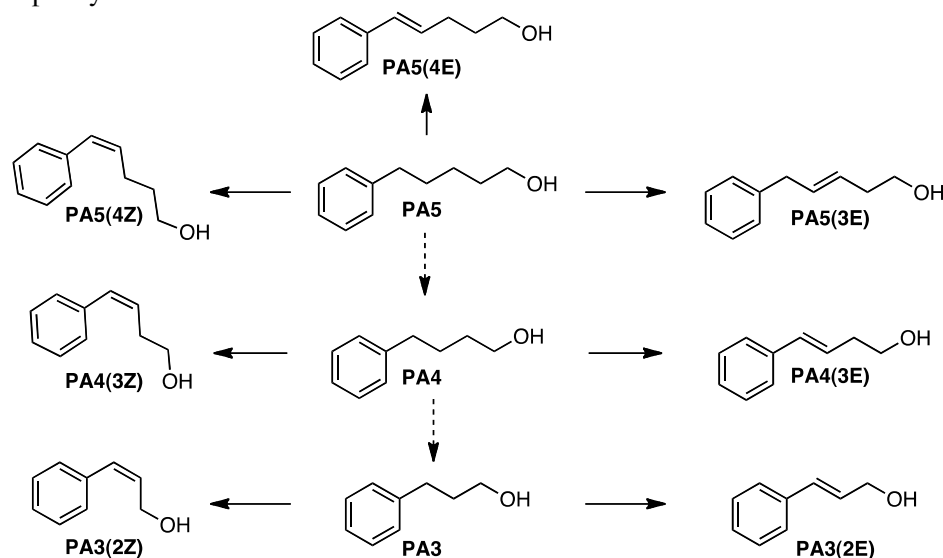
## 2.11 Observed Thermodynamics of Binding: Insertion of Internal Olefin.

**Table 2.11.1** Difference in observed ITC for the insertion of a C-C double bond into alkanols at 298 K in PBS buffer at pH 7.4.



<i>Modification</i>	$\Delta\Delta G_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )
A6→A6(2E)	0.3 ± 0.1	1.2 ± 0.1	-0.8 ± 0.2
A6→A6(2Z)	0.8 ± 0.4	2.4 ± 0.07	-1.5 ± 0.3
A6→A6(3E)	0.5 ± 0.5	2.4 ± 1.0	-1.9 ± 0.5
A6→A6(3Z)	0.6 ± 0.5	2.4 ± 1.0	-1.8 ± 0.6
A6→A6(5)	0.7 ± 0.05	0.9 ± 0.7	-0.2 ± 0.2
A7→A7(2E)	-0.3 ± 0.05	0.8 ± 0.3	-1.1 ± 0.3
A7→A7(3Z)	0.7 ± 0.1	3.4 ± 0.2	-2.7 ± 0.2
A8→A8(2E)	0.3 ± 0.1	1.9 ± 0.3	-1.6 ± 0.4
A8→A8(3Z)	0.7 ± 0.1	2.6 ± 0.2	-1.9 ± 0.2
Average*	0.5 ± 0.07	2.1 ± 0.1	-1.7 ± 0.1
(E)	0.2 ± 0.1	1.6 ± 1.2	-1.3 ± 0.2
(Z)	0.7 ± 0.05	2.7 ± 0.1	-2.0 ± 0.1

**Table 2.11.2** Difference in observed ITC for the insertion of a C-C double bond into the phenylalkanols.



<i>Modification</i>	$\Delta\Delta G_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_{obs}^{\circ}$ (kcal•mol <sup>-1</sup> )
PA3→PA3(2E)	0.0 ± 0.2	0.7 ± 0.6	-0.7 ± 0.1
PA3→PA3(2Z)	0.2 ± 0.02	3.9 ± 0.7	-3.6 ± 0.1
PA4→PA4(3E)	0.2 ± 0.02	1.7 ± 1.0	-1.6 ± 0.1
PA4→PA4(3Z)	0.1 ± 0.02	2.7 ± 0.9	-2.7 ± 0.1
PA5→PA5(3E)	0.4 ± 0.1	3.8 ± 0.8	-3.4 ± 0.6
PA5→PA5(4E)	0.02 ± 0.07	3.1 ± 0.9	-3.0 ± 0.3
PA5→PA5(4Z)	0.6 ± 0.07	3.4 ± 1.0	-2.7 ± 0.3
Average	0.2 ± 0.05	2.8 ± 0.3	-2.6 ± 0.5
(E)	0.2 ± 0.05	2.4 ± 0.4	-2.2 ± 0.2
(Z)	0.3 ± 0.02	3.3 ± 0.5	-3.0 ± 0.1

The difference in the observed thermodynamic parameters accompanying the introduction of an internal olefin between the *n*-alkanols and phenylalkanols series and each analogous member of the *n*-alkenol and phenylalkenol series are given in Tables 2.11.1 and 2.11.2, respectively. Accordingly, we observed that for the *n*-alkanols, a favorable  $T\Delta\Delta S^\circ_{\text{obs}}$  of 1.7 kcal•mol<sup>-1</sup> is compensated by an unfavorable  $\Delta\Delta H^\circ_{\text{obs}}$  of 2.1 kcal•mol<sup>-1</sup> *but to a weaker extent* than was observed for methylene removal, such that the resulting alkenols are equipotent. Hence, we do observe a favorable change in the entropy of binding upon the insertion of a C-C double bond, which is in agreement with what is expected for the lowering of ligand conformational entropy, this benefit in entropy is wiped out by a penalty in binding enthalpy that presumably arises from entropy/enthalpy compensation.

Similarly for the phenylalkanols, a favorable change in binding entropy of 2.6 kcal•mol<sup>-1</sup> accompanies the insertion of an internal olefin that is also compensated by an unfavorable change in enthalpy of 2.8 kcal•mol<sup>-1</sup> such that the resulting phenylalkenols are equipotent. Again, we observe that although unfavorable changes in  $\Delta\Delta H^\circ$  compensate favorable changes in entropy, but to a weaker extent than was observed for the removal of a methylene. The average values are also given in Tables 2.11.1 and 2.11.2.

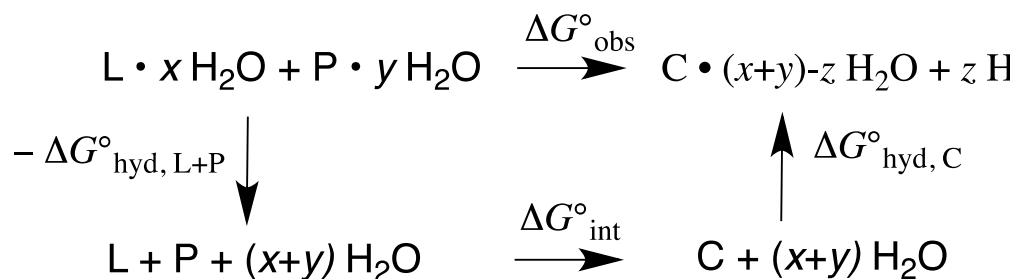
The  $\Delta\Delta G^\circ_{\text{obs}}$  for the introduction of *Z*-olefins into the *n*-alkanols and phenylalkanols is 0.5 and 0.1 kcal•mol<sup>-1</sup> more unfavorable, respectively, than that for the introduction of *E*-olefins, which for all intents and purposes is a rather insignificant difference. That larger differences in the free energies of binding between *Z*- and *E*-olefins are not observed *suggests the lack of any*

*configurational preference by the protein.* Interestingly, we observed that  $T\Delta\Delta S_{\text{obs}}^{\circ}$  for the introduction of the terminal olefin is on average more favorable than the modification of methylene removal, which is interesting. One notable exception to this trend is that the modification of **A6** to **A6(5)** is only +0.9 kcal•mol<sup>-1</sup>, which is similar to what was observed for the removal of a methylene group. These changes in thermodynamic parameters upon binding to MUP-I for the insertion of an internal olefin into the phenyl and alkylalka(e)nols depict that there is a *larger entropic benefit* upon binding for modifying the ligand with an internal olefin than for methylene removal.

There have been few systematic studies on the utility of the introduction of an internal C–C double bond as a means to preorganize a protein-binding ligand into a constrained conformation,<sup>cxxxviii,cxxxix</sup> and to the best of our knowledge there has been no known evaluation of this class of rotor restriction that includes ITC analysis. In a rather encompassing study of  $\mu$ -opioid receptors (MOR), Verdine and coworkers inserted both *E*- and *Z*-olefins into methyl linkers of several closely related, flexible mimics of a known high-affinity binder.<sup>cxl</sup> It was discovered that the ligands containing *E*-olefins generally bound better than the flexible mimics, which subsequently bound better than the ligands restricted with *Z*-olefins in the methylene linker. That this trend is not observed in our system is not all that surprising given the directional interactions that are required for favorable binding interactions to MORs, which is not the case for MUP-I since its binding interactions are dominated by non-directional vdW interactions.

*Intrinsic Analysis:*

Water plays a significant role in the thermodynamics of protein-ligand interactions, but quantifying its role is difficult. While displacing a bound water molecule during complex formation is generally regarded as having an entropically favorable consequence,<sup>exli</sup> predicting the contribution of solvent reorganization to binding thermodynamics is a formidable task because it depends upon relative changes in the dispersive interactions arising from protein and ligand desolvation<sup>exlii</sup> and upon the surface topology of the protein binding domain and the ligand.<sup>exliii</sup> Protein-ligand binding can be decomposed by way of a Born-Haber cycle (see Section 1.2) to obtain the thermodynamic parameters arising from complex formation in the absence of solvent, herein referred to as the ‘intrinsic’ thermodynamic parameters. Accordingly, a protein-ligand interaction in water may be described as follows:



where L, P, and C denote ligand, protein, and complex, respectively, and the number of water molecules hydrating each is given by  $x$ ,  $y$ , and  $z$ . The observed free energy of complexation  $\Delta G^\circ_{\text{obs}}$  is then given by the expression:

$$\Delta G^\circ_{\text{obs}} = \Delta G^\circ_{\text{int}} - \Delta G^\circ_{\text{hyd}, \text{L}+\text{P}} + \Delta G^\circ_{\text{hyd}, \text{C}} \quad (2.3)$$

where  $\Delta G_{\text{int}}^{\circ}$  is the intrinsic free energy for complex formation,  $\Delta G_{\text{hyd,L+P}}^{\circ}$  is the sum of the free energies of hydration of unbound ligand and *apo* protein, and  $\Delta G_{\text{hyd,C}}^{\circ}$  is the free energy of hydration of the protein-ligand complex. If one makes the simplifying assumption that the free energy change associated with hydration of the complex,  $\Delta G_{\text{hyd,C}}^{\circ}$ , is the same for two *closely* related ligands  $L_1$  and  $L_2$  interacting with the same protein, then the difference in the intrinsic free energies of complex formation between ligands  $L_1$  and  $L_2$ ,  $\Delta\Delta G_{\text{int},2-1}^{\circ}$ , is given by the expression:

$$\begin{aligned}\Delta\Delta G_{\text{int},2-1}^{\circ} &= \Delta G_{\text{int},2}^{\circ} - \Delta G_{\text{int},1}^{\circ} = \\ &(\Delta G_{\text{obs},2}^{\circ} - \Delta G_{\text{obs},1}^{\circ}) + (\Delta G_{\text{hyd},2}^{\circ} - \Delta G_{\text{hyd},1}^{\circ})\end{aligned}\quad (2.4)$$

Given the structural simplicity of the ligands employed in this study, it is possible to estimate their hydration free energies,  $\Delta G_{\text{hyd,L}}^{\circ}$ , and enthalpies,  $\Delta H_{\text{hyd,L}}^{\circ}$ , using group contributions that are based on compilations of experimentally determined values for hydrocarbons and monohydric alcohols at 298 K.<sup>cxliv</sup> With these values, differences in the intrinsic thermodynamic parameters of binding,  $\Delta\Delta G_{\text{int}}^{\circ}$ ,  $\Delta\Delta H_{\text{int}}^{\circ}$ , and  $T\Delta\Delta S_{\text{int}}^{\circ}$ , for methylene group removal and the introduction of an internal olefin were calculated.

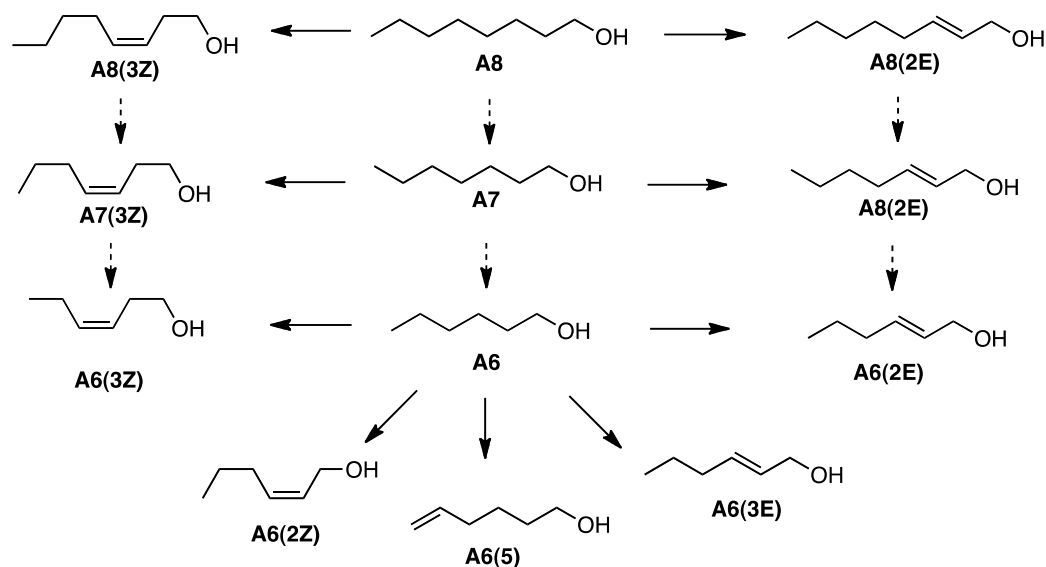
We observed that the differences in the intrinsic thermodynamic values accompanying the removal of methylene groups for the alkyl alka(e)nols are similar to differences in the observed values. Namely, the individual modifications are characterized by a favorable  $T\Delta\Delta S_{\text{int}}^{\circ}$  that is dominated by an unfavorable  $\Delta\Delta H_{\text{int}}^{\circ}$  with the result  $\Delta\Delta G_{\text{int}}^{\circ}$  for each is unfavorable. For the



phenyl alka(e)nols we see this same trend such that favorable gains in binding entropy are negated by unfavorable changes in binding enthalpy. The average values over the set of *n*-alka(e)nols and phenylalka(e)nols are given in Tables 2.11.3 and 2.11.4 respectively.

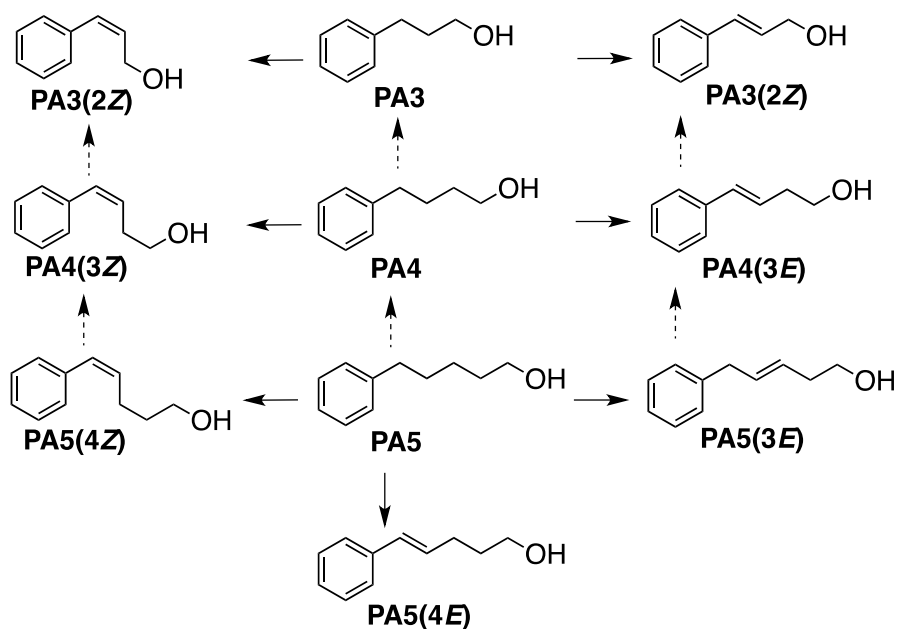
When we compare the phenyl and alka(e)nols, we find that the average  $\Delta\Delta G_{\text{int}}^{\circ}$  for the removal of a methylene group from the *n*-alka(e)nols is more unfavorable by 0.4 kcal•mol<sup>-1</sup> than for the phenyl alka(e)nols because the average  $T\Delta\Delta S_{\text{int}}^{\circ}$  of the former is less favorable by a statistically significant 1.5 kcal•mol<sup>-1</sup> than that of the latter. The average  $\Delta\Delta H_{\text{int}}^{\circ}$  for the *n*-alka(e)nols is less favorable than that for the phenylalka(e)nols is evident, but the difference is not statistically significant. Because the observed and intrinsic thermodynamic values accompanying the removal of methylene groups for the two series of ligands exhibit a similar trend, our analysis suggests that differences between the two series *do not arise from solvent reorganization* accompanying binding. Rather, they are more likely the result of dissimilarities in dispersive or vdW interactions in the different protein-ligand complexes. We consider this a significant discovery for the utility of intrinsic analysis, as our results have revealed important details regarding specific protein-ligand interactions that were before unforeseen.

**Table 2.11.3** Intrinsic difference in observed ITC for the removal of a methylene group in alkylalkanols. Work done by Richard Malham (Leeds University).



<i>Modification</i>	$\Delta\Delta G_i^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_i^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_i^\circ$ (kcal•mol <sup>-1</sup> )
<b>A8→A7</b>	<b>0.7 ± 0.9</b>	<b>1.9 ± 0.6</b>	<b>-1.2 ± 0.6</b>
<b>A7→A6</b>	<b>0.9 ± 0.7</b>	<b>2.5 ± 0.6</b>	<b>-1.6 ± 0.7</b>
<b>A8(2E)→A7(2E)</b>	<b>0.1 ± 2.2</b>	<b>0.8 ± 1.1</b>	<b>-0.7 ± 0.8</b>
<b>A7(2E)→A6(2E)</b>	<b>1.5 ± 2.2</b>	<b>3.0 ± 1.1</b>	<b>-1.5 ± 0.9</b>
<b>A8(3Z)→A7(3Z)</b>	<b>0.7 ± 0.1</b>	<b>2.7 ± 0.5</b>	<b>-2.0 ± 0.6</b>
<b>A7(3Z)→A6(3Z)</b>	<b>0.8 ± 0.1</b>	<b>1.6 ± 0.6</b>	<b>-0.8 ± 0.7</b>
<b>Average</b>	<b>0.8 ± 0.1</b>	<b>2.1 ± 0.3</b>	<b>-1.3 ± 0.3</b>

**Table 2.11.4** Intrinsic difference in observed ITC for the removal of a methylene group in phenylalkanols.

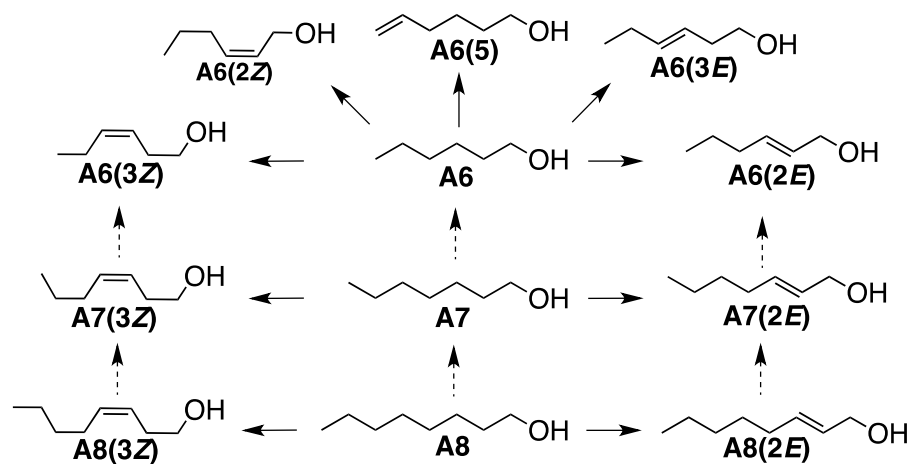


<i>Ligand</i>	$\Delta\Delta G_i^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_i^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_i^\circ$ (kcal•mol <sup>-1</sup> )
PA5→PA4	0.7 ± 0.3	3.5 ± 4.2	-2.9 ± 0.3
PA4→PA3	0.4 ± 0.2	2.5 ± 3.6	-2.2 ± 0.1
PA5(4E)→PA4(3E)	0.8 ± 0.2	2.2 ± 3.6	-1.4 ± 0.2
PA4(3E)→PA3(2E)	0.2 ± 0.2	1.5 ± 3.5	-1.3 ± 0.2
PA5(4Z)→PA4(3Z)	0.2 ± 0.2	2.9 ± 3.9	-2.8 ± 0.2
PA4(3Z)→PA3(2Z)	0.2 ± 0.2	1.5 ± 3.5	-1.3 ± 0.2
Average	0.1 ± 0.2	3.6 ± 3.3	-3.1 ± 0.2

*Intrinsic Thermodynamics of Binding: Insertion of Internal Olefin* The average values over the set of *n*-alkanols (collected by Richard Malham) and phenylalkanols are given in Tables 2.11.5 and 2.11.6. Differences in the intrinsic thermodynamic values resulting from the introduction of an internal olefin do not follow the same trend as the observed values. Whereas  $\Delta\Delta G_{\text{obs}}^{\circ}$  is relatively equipotent for most of the individual modifications where an internal olefin is introduced because a favorable  $T\Delta\Delta S_{\text{obs}}^{\circ}$  *is compensated by* an unfavorable  $\Delta\Delta H_{\text{obs}}^{\circ}$ ,  $\Delta\Delta G_{\text{int}}^{\circ}$  is favorable for all modifications because a favorable  $T\Delta\Delta S_{\text{int}}^{\circ}$  *dominates* an unfavorable  $\Delta\Delta H_{\text{int}}^{\circ}$ . Furthermore, the effect of solvent on enthalpy is minimal in that  $\Delta\Delta H_{\text{int}}^{\circ}$  is only slightly more favorable than  $\Delta\Delta H_{\text{obs}}^{\circ}$  for the insertions of an internal olefin. It is the effect of solvent on entropy that dictates the sign of the free energy differences associated with introducing an internal olefin. For the individual modifications,  $T\Delta\Delta S_{\text{int}}^{\circ}$  is more favorable than  $T\Delta\Delta S_{\text{obs}}^{\circ}$  by  $\sim 1.2 \text{ kcal}\cdot\text{mol}^{-1}$  because the entropy associated with desolvation of the saturated ligands is more favorable than that for the unsaturated analogs.

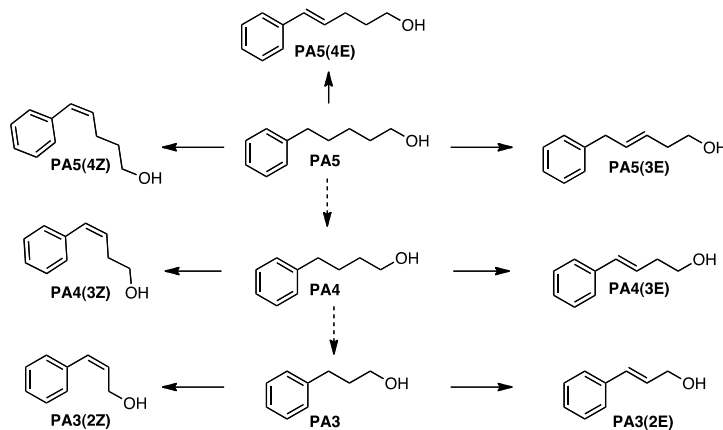
The average  $\Delta\Delta G_{\text{int}}^{\circ}$  for the introduction of an internal olefin into the *n*-alkanols is  $0.2 \text{ kcal}\cdot\text{mol}^{-1}$  less favorable than that for phenylalkanols because the average  $T\Delta\Delta S_{\text{int}}^{\circ}$  for the former is less favorable by a statistically significant  $0.9 \text{ kcal}\cdot\text{mol}^{-1}$  than the latter. Hence, the data suggests that the phenylalkanols benefit from the introduction of an internal olefin to a slightly greater extent than the *n*-alkanols due to a more favorable entropic component that may be the result of reduced conformational entropy in the unbound state. However, this benefit is negated by an enthalpic penalty.

**Table 2.11.5** Intrinsic difference in intrinsic ITC for the insertion of a C–C double bond into the alkylalkanols, done by Richard Malham



<i>Modification</i>	$\Delta\Delta G_i^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_i^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_i^\circ$ (kcal•mol <sup>-1</sup> )
<b>A6→A6(2E)</b>	<b>-0.6 ± 0.2</b>	<b>1.4 ± 0.7</b>	<b>-2.0 ± 0.8</b>
<b>A6→A6(2Z)</b>	<b>-0.3 ± 0.2</b>	<b>2.3 ± 0.6</b>	<b>-2.6 ± 0.6</b>
<b>A6→A6(3E)</b>	<b>-0.4 ± 0.2</b>	<b>2.5 ± 0.6</b>	<b>-3.0 ± 0.5</b>
<b>A6→A6(3Z)</b>	<b>-0.3 ± 0.2</b>	<b>2.6 ± 0.6</b>	<b>-2.9 ± 0.7</b>
<b>A6→A6(5)</b>	<b>0.02 ± 0.2</b>	<b>1.1 ± 0.5</b>	<b>-1.1 ± 0.7</b>
<b>A7→A7(2E)</b>	<b>-1.2 ± 0.5</b>	<b>0.9 ± 1.1</b>	<b>-2.2 ± 0.6</b>
<b>A7→A7(3Z)</b>	<b>-0.2 ± 0.1</b>	<b>3.6 ± 0.6</b>	<b>-3.8 ± 0.7</b>
<b>A8→A8(2E)</b>	<b>-0.6 ± 0.3</b>	<b>2.1 ± 0.6</b>	<b>-2.7 ± 0.6</b>
<b>A8→A8(3Z)</b>	<b>-0.2 ± 0.3</b>	<b>2.8 ± 0.6</b>	<b>-3.0 ± 0.6</b>
<b>Average*</b>	<b>-0.5 ± 0.07</b>	<b>2.3 ± 0.07</b>	<b>-2.8 ± 0.2</b>
<b>(E)</b>	<b>-0.3 ± 0.1</b>	<b>1.7 ± 0.4</b>	<b>-2.4 ± 0.4</b>
<b>(Z)</b>	<b>-0.2 ± 0.4</b>	<b>2.8 ± 0.3</b>	<b>-3.1 ± 0.3</b>

**Table 2.11.6.** Intrinsic difference in intrinsic ITC for the insertion of a C–C double bond into the phenylalkanols



<i>Modification</i>	$\Delta\Delta G_i^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta\Delta H_i^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta\Delta S_i^\circ$ (kcal•mol <sup>-1</sup> )
PA3→PA3(2E)	-0.9 ± 0.2	1.0 ± 0.8	-1.9 ± 0.5
PA3→PA3(2Z)	-0.7 ± 0.2	4.0 ± 0.8	-4.7 ± 0.5
PA4→PA4(3E)	-0.7 ± 0.2	2.0 ± 1.2	-2.7 ± 0.5
PA4→PA4(3Z)	-0.8 ± 0.2	3.0 ± 1.1	-3.8 ± 0.5
PA5→PA5(3E)	-0.5 ± 0.3	4.0 ± 1.0	-4.5 ± 0.8
PA5→PA5(4E)	-0.9 ± 0.2	3.2 ± 1.1	-4.1 ± 0.6
PA5→PA5(4Z)	-0.3 ± 0.2	3.6 ± 1.1	-3.9 ± 0.6
Average	-0.7 ± 0.07	3.0 ± 0.4	-3.7 ± 0.2
(E)	-0.8 ± 0.1	2.6 ± 0.5	-3.3 ± 0.3
(Z)	-0.6 ± 0.1	3.5 ± 0.6	-4.1 ± 0.2

We noticed that the introduction of an internal olefin into the *n*-alkanols and phenylalkanols gave intrinsic entropies that were consistently two-fold larger in magnitude than those observed for the removal of a methylene unit. It should be mentioned that this result, taken together with the previously discussed result of inserting a terminal olefin into **A6** suggests that an internal olefin yields more favorable intrinsic entropies than either introducing a terminal olefin or removal of a methylene group.

It should be stated that the solvation parameters for calculating the intrinsic values do not depend on location or geometry of the olefin. In general, we found that  $\Delta\Delta G_{\text{int}}^{\circ}$  for the introduction of *E*-olefins into both the *n*-alkanol and phenylalkanol series is 0.5 and 0.2 kcal•mol<sup>-1</sup> more favorable, respectively, than that for the introduction of *Z*-olefins because  $\Delta\Delta H_{\text{int}}^{\circ}$  for the introduction of *Z*-olefins is less unfavorable. Since larger differences in the free energies of binding between *Z*- and *E*-olefins are not observed, there appears to be no configurational preference by the protein.

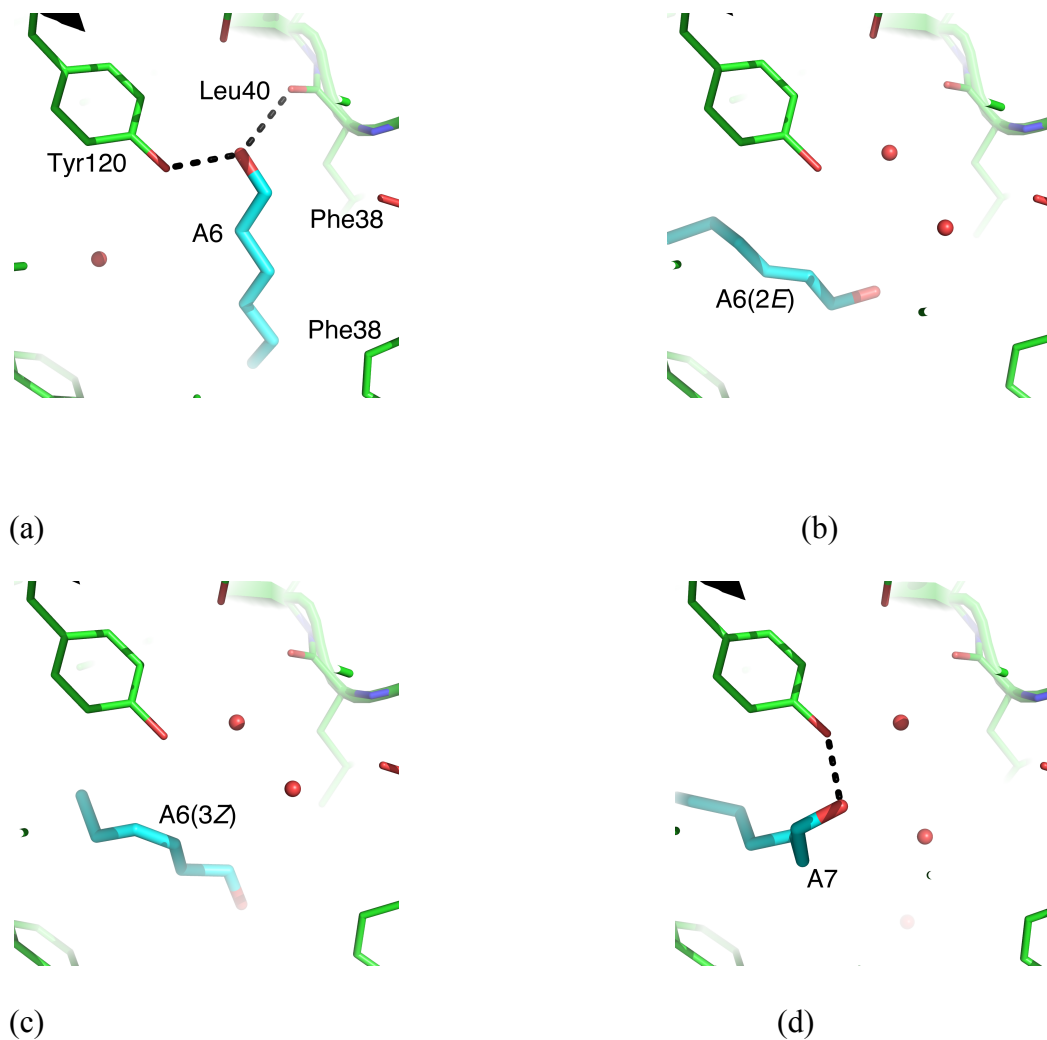
## 2.12 Crystal Structures

*Crystallographic Analysis.* X-ray crystal structures of MUP-I in complex with alkyl alcohols **A6**, **A6(3Z)**, **A6(2E)**, **A7**, **A7(3Z)**, **A7(2E)**, **A8**, **A8(3Z)**, and **A8(2E)** were obtained to a resolution of <1.5 Å by C. Lee at Leeds University. The structures of the complexes with aryl alcohols **PA3**, **PA3(2E)**, **PA4**, **PA4(3E)**, **PA5**, and **PA5(3E)**, were obtained to a resolution of <2.0 Å by research scientist Dr. John Clements. Backbone atoms belonging to the protein in the complexes of the *n*-alka(e)nols align closely, with pair wise RMSDs <0.2 Å. However, the position

and conformation of the bound ligands in the pocket vary considerably from one complex to another. For example, while **A6** and **A7** both bind such that their hydroxyl oxygen atoms make direct polar contacts with the hydroxyl oxygen atom of Tyr120, the alkyl chains belonging to these ligands extend in different directions within the pocket, and their terminal methyl carbon atoms of the terminal methyl groups occupy positions that are  $>9 \text{ \AA}$  from one another in aligned complexes. This is not an unusual observation given that the lipophilic molecule 2-sec-butyl-4,5-dihydroxythiazole was found via X-ray crystallography to bind to MUP-I in different poses.<sup>cxlv</sup> Moreover, it was previously shown by the Homans groups that 1-pentanol bound to MUP-I with different binding modes, each having partial occupancy.

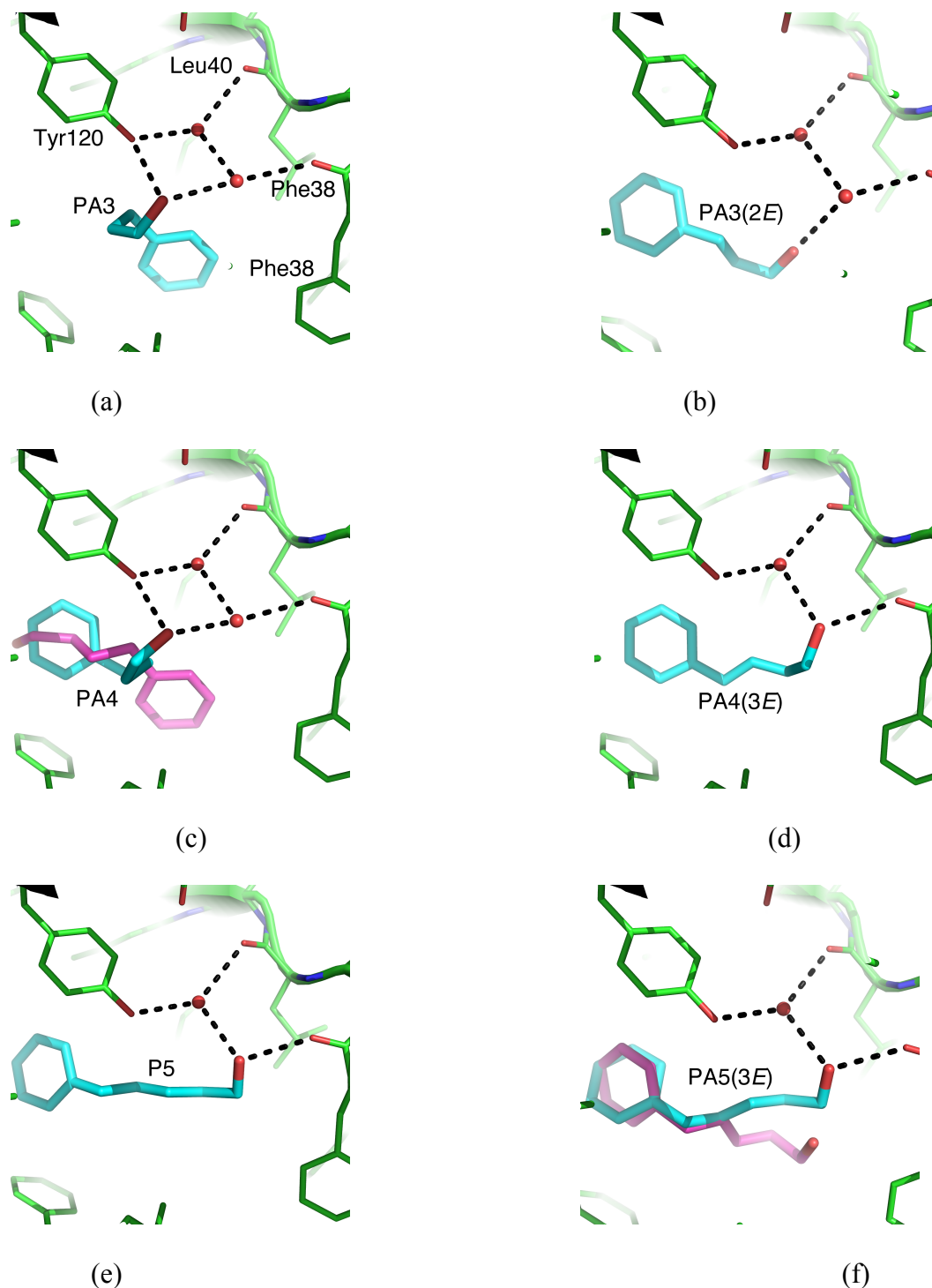


**Figure 2.12.1** Complexes of MUP-I with the *n*-alka(e)nols obtained at Leeds University. Oxygen and nitrogen atoms are colored red and blue, respectively.<sup>a</sup>



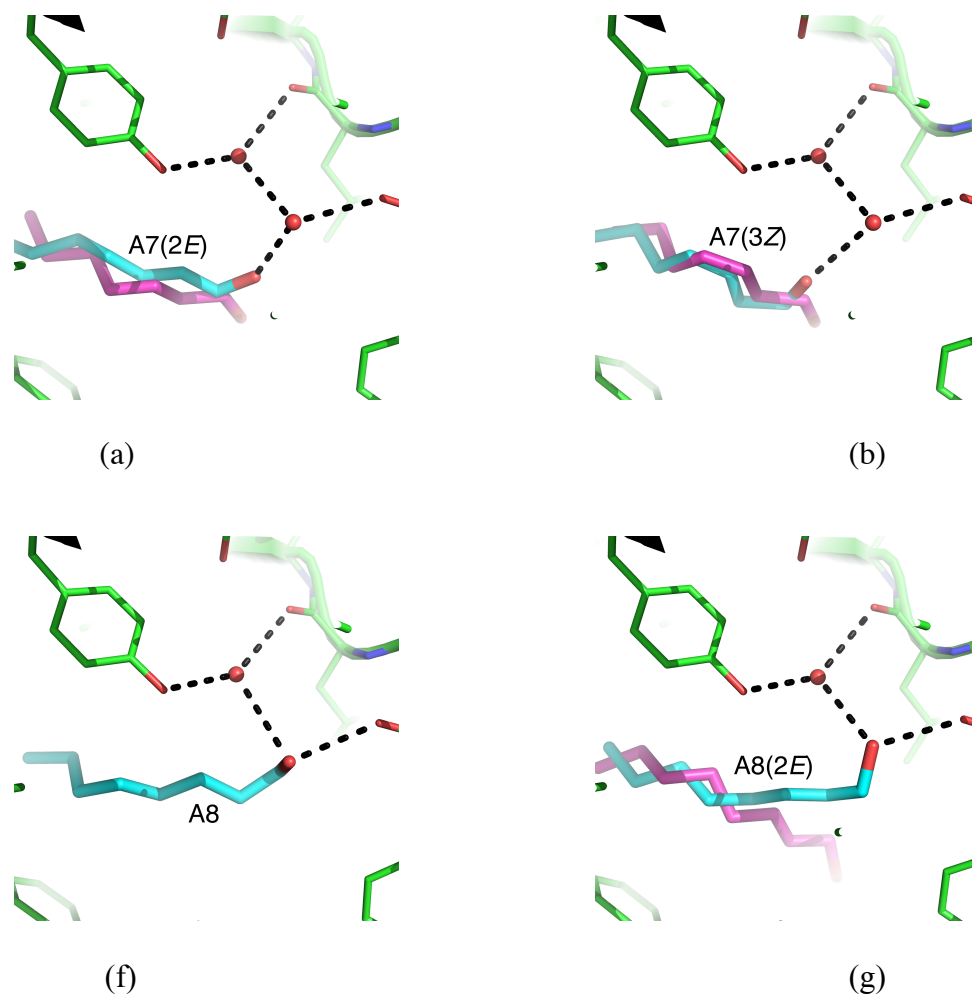
<sup>a</sup>Carbon atoms belonging to the protein (line, cartoon) are colored green while those belonging to the ligands (sticks) are colored cyan or magenta, the latter shown for complexes in which the bound ligands occupy more than one pose. Protein-ligand direct contacts and those mediated by water molecules (red spheres) are indicated with dashed, black lines. (a) Complex with A6. (b) Complex with A6(2E). (c) Complex with A6(2Z). (d) Complex with A7.

**Figure 2.12.2** Complexes of MUP-I with the *n*-alka(e)nols.<sup>a</sup>



<sup>a</sup>Oxygen and nitrogen atoms are colored red and blue, respectively. Carbon atoms belonging to the protein (line, cartoon) are colored green while those belonging to the ligands (sticks) are colored cyan or magenta, the latter shown for complexes in which the bound ligands occupy more than one pose. Protein-ligand direct contacts and those mediated by water molecules (red spheres) are indicated with dashed, black lines. (a) Complex with PA3. (b) Complex with PA3(2E). (c) Complex with PA4. (d) Complex with PA4(3E). (e) Complex with P5. (f) Complex with PA5(3E).

**Figure 2.12.3** Complexes of MUP-I with the *n*-alka(e)nols.<sup>a</sup>



<sup>a</sup>Oxygen and nitrogen atoms are colored red and blue, respectively. Carbon atoms belonging to the protein (line, cartoon) are colored green while those belonging to the ligands (sticks) are colored cyan or magenta, the latter shown for complexes in which the bound ligands occupy more than one pose. Protein-ligand direct contacts and those mediated by water molecules (red spheres) are indicated with dashed, black lines. (e) Complex with **A7(2E)**. (f) Complex with **A7(3Z)**. (g) Complex with **A8**. (h) Complex with **A8(2E)**.

While some of the ligands in the complexes participate in direct and/or water-mediated polar contacts to MUP-I, others do not (Figure 2.13.2-2.13.4), suggesting that maintaining a conserved protein-ligand polar contact network is not a prerequisite for the binding of ligands to this protein. For example, *n*-alkenols **A6(2E)** and **A6(3Z)** make no direct or water-mediated contacts to the protein in their respective complexes (Figure 2.13.2). It is clear that certain residues belonging to the protein do recur frequently when analyzing polar networks in the different complexes, and these are Phe38, Leu40, and Tyr120 (Figure 2.13.1 a), which have polar interactions with the ligands either directly or through water-mediated contacts in several of the complexes, the former two via their backbone carbonyl oxygen atoms.

In those complexes for which ligands bind in two different poses (Figure 2.13.2 d, g and Figure 2.13.3 a, b, g), the OH group of the ligand in one pose makes more polar contacts with Tyr120, but in the other pose the ligand makes no such contacts. Inspection of the images also leads us to conclude that longer *n*-alka(e)nols tend to make more polar contacts with Tyr120 and water networks than shorter ones due to more extensive H<sub>2</sub>O networks. This observation would seem to run counter to the previous assumption that complexes of related ligands have similar thermodynamic parameters of hydration; However, upon examining closely related ligands, *e.g.* those differing by a single methylene unit or internal double bond (Figure 2.12.3 a and g), the assumption appears to still be valid.

As observed in Figure 2.12.3 c, polar protein networks with **PA4** involving direct and water-mediated contacts to Phe38, Leu40, and Tyr120 of MUP-I were also

observed for the phenylalka(e)nols. Once again, in the complexes for which ligands bind in two different poses, the ligand in one makes more polar contacts with the protein, but in the other pose it does not. Unlike the *n*-alka(e)nols, however, the networks of polar contacts do not appear to be dependent on the size of the ligand *d* (Figure 2.13.4), and similar contact networks were observed for all ligands in the phenylalka(e)nol series. Given the size of MUP-I binding pocket, it is interesting that, with the exception of **A6**, no more than two crystallographic water molecules were observed in the pocket of the complexes with *n*-alka(e)nols, phenylalka(e)nols, or the *apo* structure,<sup>cxlvb</sup> supporting the hypothesis that the pocket may be dewetted.<sup>cxlvi</sup>

*Assessing the assumption that hydration differences of protein complexes for two comparable ligands are minimal.* It was crucial for us to scrutinize the role of water, because assuming solvation of the protein complex is the same for different ligands is necessary to calculate our intrinsic binding data. Accordingly, the number of crystallographic water molecules observed in the binding pocket of the complexes were tallied and differences in this number were recorded for ligands based upon methylene group removal and introduction of an internal olefin (Table 2.12.1). For both the *n*-alka(e)nols and the phenylalka(e)nols, methylene group removal is accompanied by an *average increase* of 0.50 crystallographic water molecules, whereas the introduction of an internal olefin is accompanied by an *average decrease* of 0.22 water molecules (Table 2.12.2). The variations in hydration of the protein are thus evident, but these are rather small. Larger ligands tend to displace a greater number of water molecules, but there is no correlation between differences in

the number of crystallographic water molecules in the pocket and differences in ligand Connolly surface area.<sup>cxlvii</sup>

We also examined average all-atom B factors, which are thought to be a barometer of relative protein dynamics.<sup>cxlviii</sup> It is postulated that larger B-factors are an indication of greater thermal motion in the complexes for a series of varying protein-ligand bound structures. However, it has been questioned whether B-factors can be accurately correlated with thermodynamic parameters such as entropy and enthalpy. Specifically, in a series of Grb2 SH2 bound ligands wherein the B-factors were tallied, it was found that the B-factors were generally smaller for the ligands that bound with more favorable entropies compared to a set of constrained ligands that bound with more favorable enthalpies.<sup>lxxvi</sup> For this current work, we investigated All-Atom B-factors and found no significant correlations between the thermodynamic parameters of the phenyl and alkylalka(e)nols. (table 2.12.1).

**Table 2.12.1.** Structural features of the complexes.<sup>a</sup>

Complex	Average All-Atom B-factor (Å <sup>2</sup> ) <sup>[a]</sup>		No. bound	Ligand CSA	
	Protein	Pocket <sup>[b]</sup>	Ligand <sup>[c]</sup>	Water <sup>[d]</sup>	(Å <sup>2</sup> ) <sup>[c],[e]</sup>
Alka(e)nols					
<b>A6</b>	25.0	17.2	31.8	3	143.3
<b>A6(2E)</b>	20.2	14.8	37.6	2	137.7
<b>A6(3Z)</b>	18.7	14.7	24.2	2	136.4
<b>A7</b>	22.8	14.7	34.8	3	158.1
<b>A7(2E)</b>	17.9	13.6	22.4	2	156.7
			23.2		155.2
<b>A7(3Z)</b>	19.0	14.9	25.3	2	152.4
			25.1		151.1
<b>A8</b>	24.5	17.6	30.9	1	176.4
<b>A8(2E)</b>	18.6	13.8	22.0	1	174.5
			24.5		172.9
<b>A8(3Z)</b>	17.9	13.7	23.1	2	169.8
			23.0		166.7
Phenylalka(e)nols					
<b>PA3</b>	22.4	17.1	29.3	2	158.0
<b>PA3(2E)</b>	22.4	16.6	39.0	2	153.1
<b>P4</b>	24.3	17.6	20.7	2	175.7
			32.9		177.5
<b>PA4(3E)</b>	22.1	15.1	34.6	1	196.9
<b>PA5</b>	21.4	16.1	22.2	1	196.9
<b>PA5(3E)</b>	23.6	18.0	20.7	1	191.1
			22.3		189.4

[a] Average all-atom B-factor data were adjusted according to the method recommended by Ringe *et al.*<sup>cxlix</sup> for comparing data obtained from different crystals [b]. Residues comprising the binding pocket were defined as those that make at least one contact to the ligands in at least one of the complexes. These are Leu24, Phe38, Leu40, Leu42, Ile45, Leu54, Phe56, Phe90, Ile92, Leu101, Leu105, Leu116, and Tyr120. [c] In the case of multiple binding modes, average all-atom B-factor and CSA data are reported for each ligand pose. [d] Number of crystallographic water molecules found in the binding pocket. [e] Connolly surface area.<sup>cl</sup>

**Table 2.12.2.** Change in number of bound waters averaged per modification.

$\Delta$ No. bound		$\Delta$ No. bound	
<i>Modification</i>	<i>H<sub>2</sub>O</i>	<i>Modification</i>	<i>H<sub>2</sub>O</i>
<b>A6→A6(2E)</b>	<b>-1</b>	<b>A8→A7</b>	<b>2</b>
<b>A6→A6(3Z)</b>	<b>-1</b>	<b>A8(3Z)→A7(3Z)</b>	<b>0</b>
<b>A7→A7(2E)</b>	<b>-1</b>	<b>A8(2E)→A7(3E)</b>	<b>1</b>
<b>A7→A7(3Z)</b>	<b>-1</b>	<b>A7→A6</b>	<b>0</b>
<b>A8→A8(2E)</b>	<b>0</b>	<b>A7(3Z)→A6(3Z)</b>	<b>0</b>
<b>A8→A8(3Z)</b>	<b>1</b>	<b>A7(2E)→A6(2E)</b>	<b>0</b>
<b>PA3→PA3(2E)</b>	<b>0</b>	<b>PA5→PA4</b>	<b>1</b>
<b>PA4→PA4(3E)</b>	<b>1</b>	<b>PA5(3E)→PA4(3E)</b>	<b>0</b>
<b>PA5→PA5(3E)</b>	<b>0</b>	<b>PA4→PA3</b>	<b>0</b>
<b>Average</b>	<b>-0.22</b>	<b>PA4(3E)→PA3(2E)</b>	<b>1</b>
		<b>Average</b>	<b>0.50</b>

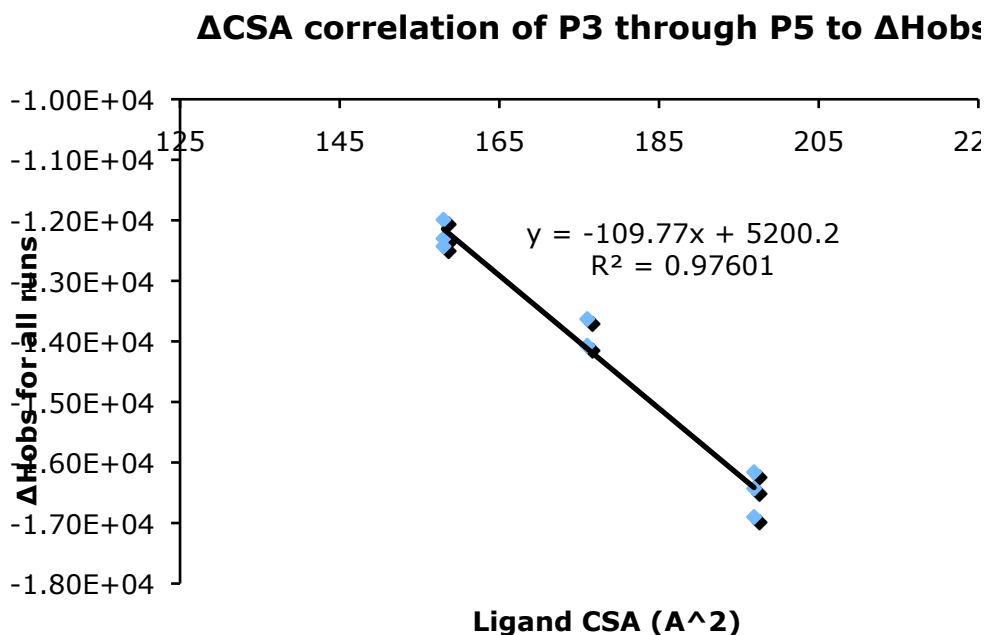
We *do* observe that CSA correlates strongly with  $\Delta H_{\text{obs}}^{\circ}$  for the Homans *n*-alkanols **A6**, **A7**, and **A8** ( $R^2 = 0.96$ ) and our phenylalkanols **PA3**, **PA4** and **PA5** ( $R^2 = 0.98$ ) with slopes of -0.08 and -0.11 kcal•mol<sup>-1</sup>Å<sup>-2</sup>, respectively (graph 2.12.1). These values closely match those obtained for the correlation of non-polar CSA and  $\Delta H_{\text{obs}}^{\circ}$  for the binding of pYXN tripeptides to Grb2 SH2 (-0.11 kcal•mol<sup>-1</sup>Å<sup>-2</sup>).<sup>cli</sup> We also observed a correlation, albeit weaker, between the number of protein-ligand



vdW contacts and  $\Delta H_{\text{obs}}^{\circ}$  for the phenylalkanols **PA3**, **PA4** and **PA5** ( $R^2 = 0.62$ ) and phenylalkenols **PA3(2E)**, **PA4(3E)** and **PA5(3E)** ( $R^2 = 0.72$ ). Since there is a correlation between the burial of  $\Delta\text{CSA}$  with an increase in tallied vdW contacts in the binding pocket, one may assume in this case that the binding is driven in part by the propensity of the ligand to make favorable vdW interactions. Since the binding pocket of the apo structure is dewetted,<sup>cxxxv</sup> then the amount of dispersive interactions gained during the protein-ligand association over-compensate for interactions that are lost between the protein and solvent.

Lastly, in an effort to elucidate the effect of changes in the number of bound water molecules on the thermodynamics of binding, the crystallographic datasets were examined for each modification type (across both observed and intrinsic data) and separated such that the individual modifications (i.e. rotor removal and rotor restriction) in which changes in the hydration of the bound complex were observed upon the testing the modified ligand (that is to say, methylene removed or olefin inserted) were placed in one subset while the individual modifications for which no changes in hydration of the bound complex were observed were placed in another subset. The two sets were compared and the average observed thermodynamic parameters associated with each were not found to be statistically different, suggesting that there are no significant differences in solvation parameters of the protein complexes over a set of *closely related* molecules.

**Graph 2.12.1.** Graph depicting the relationship between Ligand CSA and  $\Delta H_{\text{obs}}^\circ$



## 2.13 Summary and Conclusion

We studied the thermodynamic and structural effects of modifying alkyl chains of phenylalka(e)nols that bind to MUP-I by removing a rotor via deletion of a methylene unit and restriction of a rotor via the installation of an internal olefin. Our collaborators from the Homans group at Leeds University did the analogous studies on a series of *n*-alka(e)nols.

We addressed several critical questions regarding protein-ligand interactions. These included: (1) a systematic analysis of the thermodynamic effects of restricting a rotor by introducing an internal olefin, (2) unveiling the “solvent

free” or *intrinsic* binding analysis to ascertain pertinent solute-solute interactions, (3) studying the “equivalence” of rotor removal by comparing results of restricting rotors via an internal olefin vs. methylene deletion, and (4) structure-thermodynamic between thermodynamic parameters of binding such as entropy and enthalpy, and structural characteristics in the pocket such as dispersive interactions and buried Connolly surface area..

In general, we observed that for the *n*-alka(e)nols the insertion of an internal olefin gives significantly more favorable binding entropies than does methylene removal, yet enthalpy-entropy compensation leads to nearly equipotent binding energetics for the observable data. Similar thermodynamic trends were observed for the phenylalka(e)nols such that a favorable  $T\Delta\Delta S_{\text{obs}}^{\circ}$  is offset by an unfavorable  $\Delta\Delta H_{\text{obs}}^{\circ}$  so that there are no significant gains in  $\Delta\Delta G_{\text{obs}}^{\circ}$  for either removal of a methylene and insertion of an internal olefin. No significant thermodynamic distinction was observed between the introduction of *E*- and *Z*-olefins, implying that the MUP-I binding pocket has no preference for one over the other.

To develop a clearer understanding of the differences that we observed between the thermodynamics of binding of the phenylalkanols and alkylalkanols, we calculated the “solvent free” or intrinsic thermodynamics of binding to determine if differences arose from the solvent or from dispersive interactions. We employed an intrinsic analysis approach to this protein system to decompose ligand binding, which was possible because all hydration parameters of these ligands could be calculated. The results of these calculations indicated that the

*intrinsic* energetics associated with methylene group removal do not differ markedly from the observed, suggesting that the thermodynamic signature arises from dispersive interactions between the protein and the ligands and not solvent reorganization.

In contrast, the intrinsic free energies of introducing an internal olefin into the *n*-alkanols and phenylalkanols differ markedly from the observed data. The intrinsic affinities are more favorable than the observed affinities because a favorable  $T\Delta\Delta S_{\text{int}}^{\circ}$  *dominates* an unfavorable  $\Delta\Delta H_{\text{int}}^{\circ}$ . There is little difference between the observed and intrinsic enthalpies, so it suggests that the effect of solvent reorganization on the entropic component likely accounts for the difference. We can thus deduce from these calculations that the enthalpy/entropy compensation that we observe in the thermodynamic data is in part due to solvent reorganization. Since the origin of enthalpy/entropy compensation has long been sought after, our observation that solvent reorganization plays a role in this phenomenon is a significant discovery that may lead to the better understanding of solvent reorganization in protein-ligand interactions.

Furthermore, using our intrinsic analysis we also discovered that the intrinsic entropies of inserting an internal olefin are nearly double that of removing a methylene groups. This latter discovery is a potentially significant development for scoring functions that are utilized in the prediction of ligand binding affinities. These scoring functions currently assign equal and unbiased entropic penalties per rotor restricted, yet our results suggest that there is at least a

correlation with how a ligand's conformation is constrained and the binding entropy gained from rotor restriction.

Analysis of the X-ray crystallographic structures of complexes with the *n*-alka(e)nols and phenylalka(e)nols maintained any conserved polar and/or water-mediated networks, and ligands were frequently shown to bind in multiple poses. Different ligands may adopt different binding poses relative to the pocket, and the same ligand may adopt different poses. For both the *n*-alka(e)nols and the phenylalka(e)nols, methylene group removal is generally accompanied by an *increase* in the number of interfacial crystallographic water molecules whereas the introduction of an internal olefin is accompanied by a *decrease* in the number of such water molecules. However, no meaningful correlations could be found between differences in the number of crystallographic water molecules in the pocket with differences in our structural and thermodynamic parameters. Yet, we did find a strong correlation between  $\Delta G_{\text{obs}}^{\circ}$  and ligand CSA. Thus, this finding further supports the thought that MUP-I binding events are driven by favorable dispersive interactions made between the ligand and the protein in the binding pocket.

For future studies, it would be important to test more examples where a terminal alkene is bound to MUP-I to further probe the trend that we observed where upon the terminal olefin bound with significant less favorable binding entropy than the other corresponding phenyl and alkylalka(e)nols of the data set. This dataset alone could shed more light onto the favorable binding entropies that were observed upon the insertion of an internal olefin in to these MUP-I binding alkanols.

## **Chapter 3     Studies of Increasing Ligand Hydrophobicity to a Known Protein Binder**

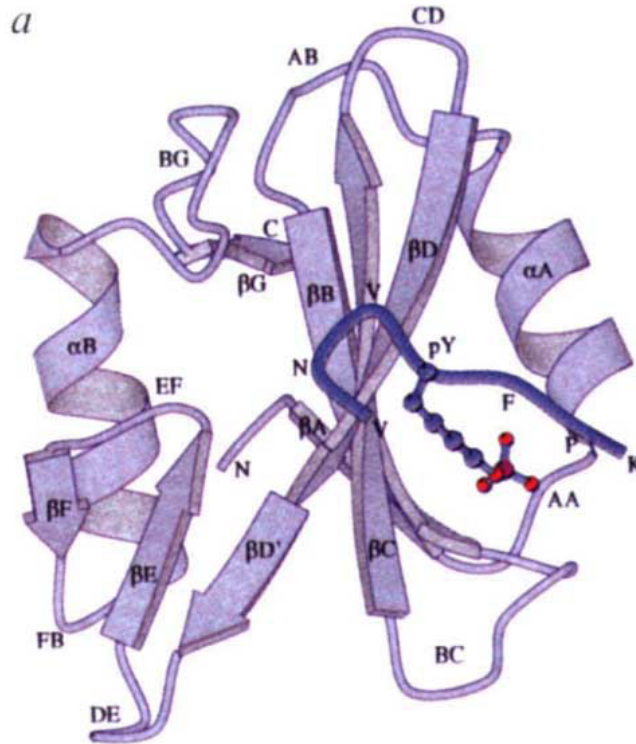
### **3.1     Introduction**

It is generally accepted in drug design that one may increase potency by adding nonpolar substituents to a drug lead.<sup>xi</sup> The rationale for this strategy is based upon a belief that the release of water molecules into the surrounding solvent during ligand desolvation is entropically favorable, in accord with the hydrophobic effect.<sup>clii,cliii</sup> However, there are few studies of such “entropy-oriented” modifications where entropy and enthalpy of binding are determined, and even fewer that correlate the binding thermodynamics with structural changes in the bound complex.<sup>cliv,clv</sup> Therefore, as part of our longstanding interest in correlating structure with binding thermodynamics of protein-ligand interactions, we became interested in the effects of adding nonpolar substituents to protein-binding ligands. One potential model system that looked intriguing to us was a class of peptidic inhibitors of the Grb2 SH2 domain that were previously pursued by the Novartis research group as potential anti-cancer therapeutics.<sup>clvi,clvii</sup>

### **3.2     Grb2 Adaptor Protein**

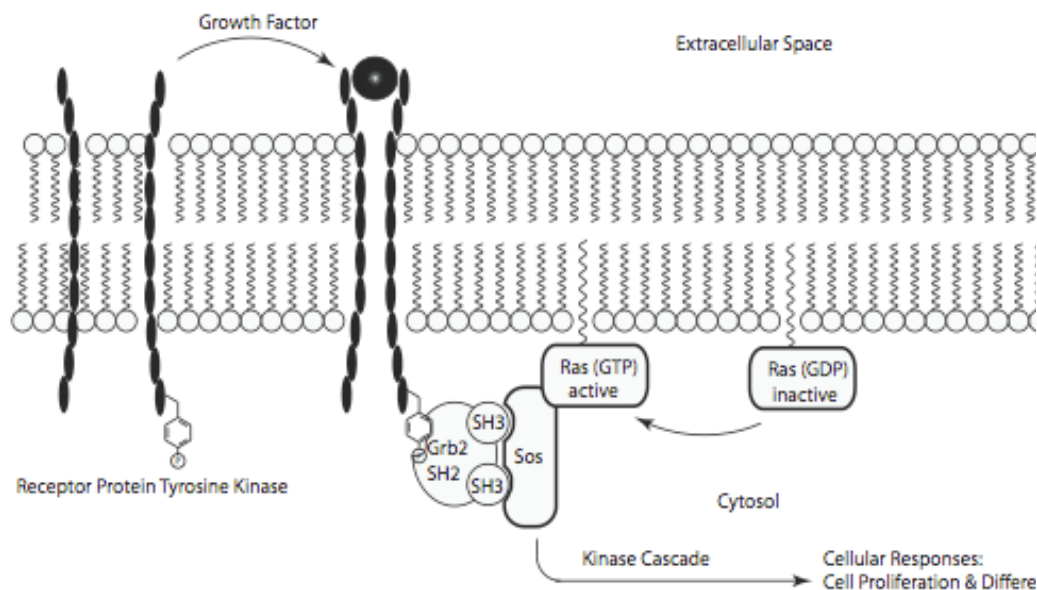
Growth receptor bound protein (Grb2) is a 25,000 KDa protein whose purpose in the mammalian cell is to function as an adaptor protein by coupling signaling receptors with effector proteins.<sup>clviii,clix</sup> Grb2 consists of one SH2 domain flanked by two SH3 domains as shown in Figure 3.2.1 bound to a peptide. This SH2 domain is known to have

a high affinity for phosphotyrosyl peptides with the sequence pYXN, where X is usually an apolar side chain. The SH3 domains are responsible for forming a cellular complex with the proline rich region of the *Son of sevenless* (Sos) effector protein.



SH2 domains have in stimulating cellular growth, selective and potent inhibitors of this domain have been proposed as viable strategy for developing anti cancer therapeutics.<sup>clxii</sup>

**Figure 3.2.2** Biological role of Grb2.

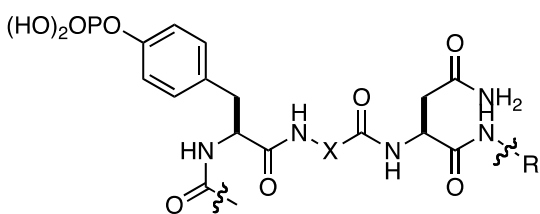


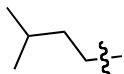
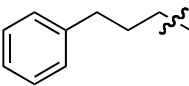
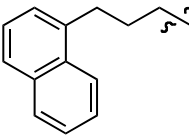
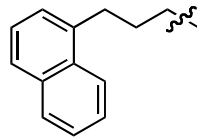
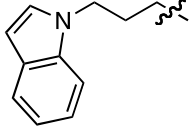
### 3.3 Experiment Design

Significant increases in ligand potency were previously observed for the pY+3 derived ligands relative to the corresponding analogs having a free C-termini shown in Figure 3.3.1. The Novartis group designed these peptides so that they could target an extended hydrophobic patch of the Grb2 SH2 domain. They postulated that they might be able to increase the potency of their initial ligands by attaching lipophilic groups to the pY+3 residue of various tripeptides. Although this strategy was shown to be successful, as a variety of the pY+3 ligands shown in figure 3.3.1 are more potent than the parent compound lacking the hydrophobic group (**1.74**), it was supported only by modeling studies (Figure 3.3.2) and neither X-ray analysis nor thermodynamic data were obtained.<sup>clvii</sup>

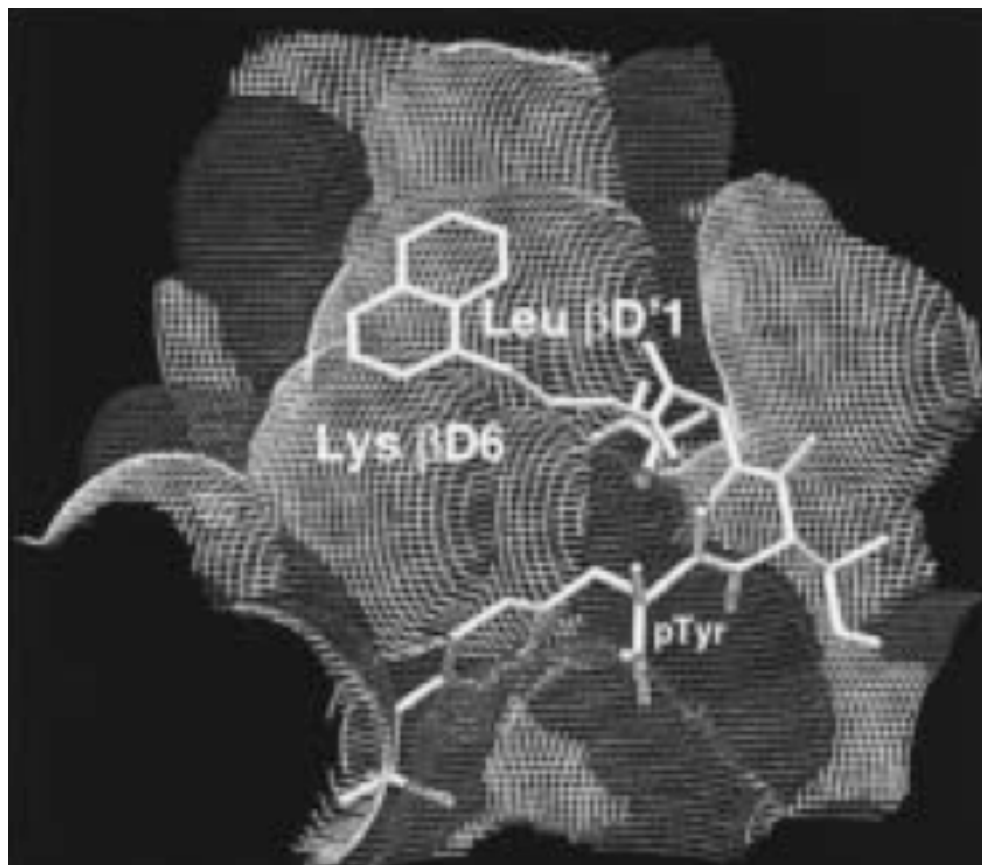


**Figure 3.3.1** Novartis Ligands containing increase apolar surface area.



compd	R =	X =	IC <sub>50</sub> (μM)
1.69		Ile	3.2 ± 0.1
1.70		Ile	2.3 ± 0.1
1.71		Ile	0.33 ± 0.1
1.72		Ac <sub>6</sub> C	0.047 ± 0.1
1.73		Ac <sub>6</sub> C	0.0012 ± 0.1
1.74	H	Ac <sub>6</sub> C	0.21 ± 0.1

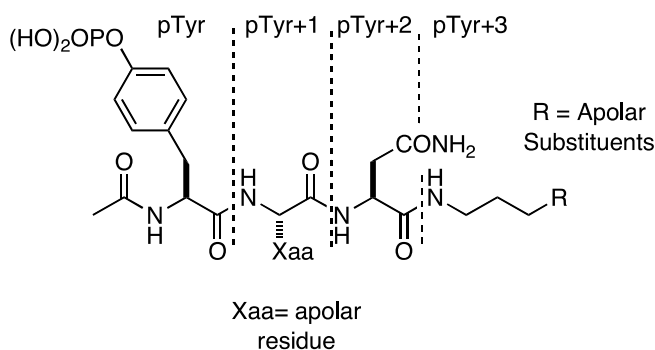
**Figure 3.3.2** Known model of pY+3 substituents interacting with an extended hydrophobic patch of the Grb2 SH2 domain.



Accordingly, we sought to *design* and *synthesize* ligands inspired by the established precedent shown in Figure 3.2.1 so that we could *study* the thermodynamic and structural effects of adding hydrophobicity in protein-ligand interactions utilizing ITC and X-ray crystallography. Given the importance of nonpolar interactions in drug discovery, we were interested in uncovering the origins of this phenomenon so that we could develop the fundamental knowledge to understand how this motif in drug design effects protein-ligand interactions. The ligands were derived from the tripeptide framework pYXN (where X is a non polar residue) containing apolar substituents that were linked to the

pY+3 position with a three carbon chain. The framework of these ligands is shown in Figure 3.3.3.

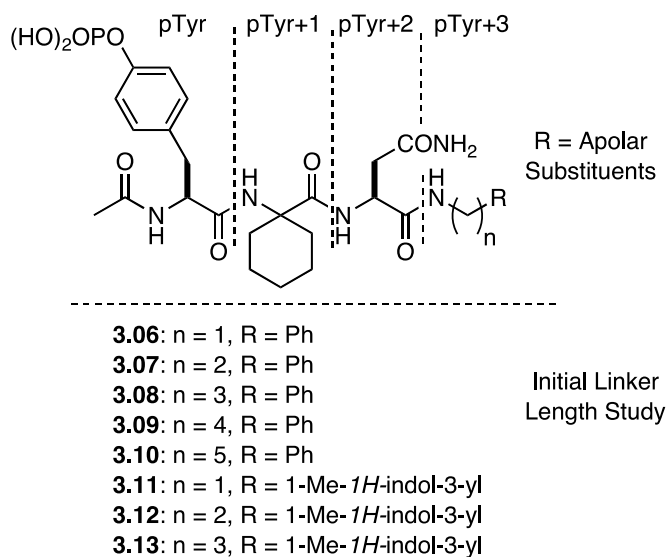
**Figure 3.3.3** Ligand framework for studies on hydrophobicity.



### 3.4 Investigation of an Appropriate Carbon Linker

Before we initiated our studies, we first had to synthesize and test ligands to determine the optimal number of methylene groups to use as a linker with the terminal groups at the pY+3 site. Although previous studies used molecular modeling to determine that  $n = 3$  is the optimal number of methylene groups to achieve favorable interactions (as shown in Figure 3.3.2),<sup>clviii</sup> it was imperative to determine this experimentally since so much of our design depended on making favorable vdW interactions with the binding pocket. Therefore, ligands of the type shown in Figure 3.4.1 where  $n$  was varied from one to five and where R = phenyl **3.06–3.09** and indolyl **3.10–3.13** were first synthesized and tested. Dr. Bo Cheng synthesized and tested the indolyl ligands **3.10–3.13**.

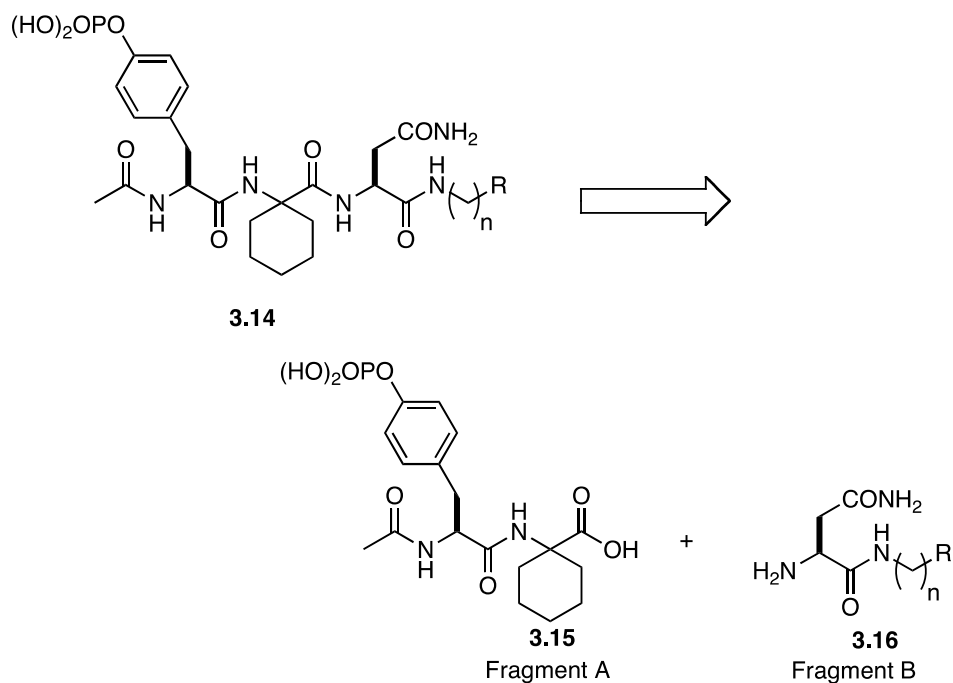
**Figure 3.4.1** Tripeptide skeleton for the hydrophobicity studies.



### 3.4 Ligand Synthesis

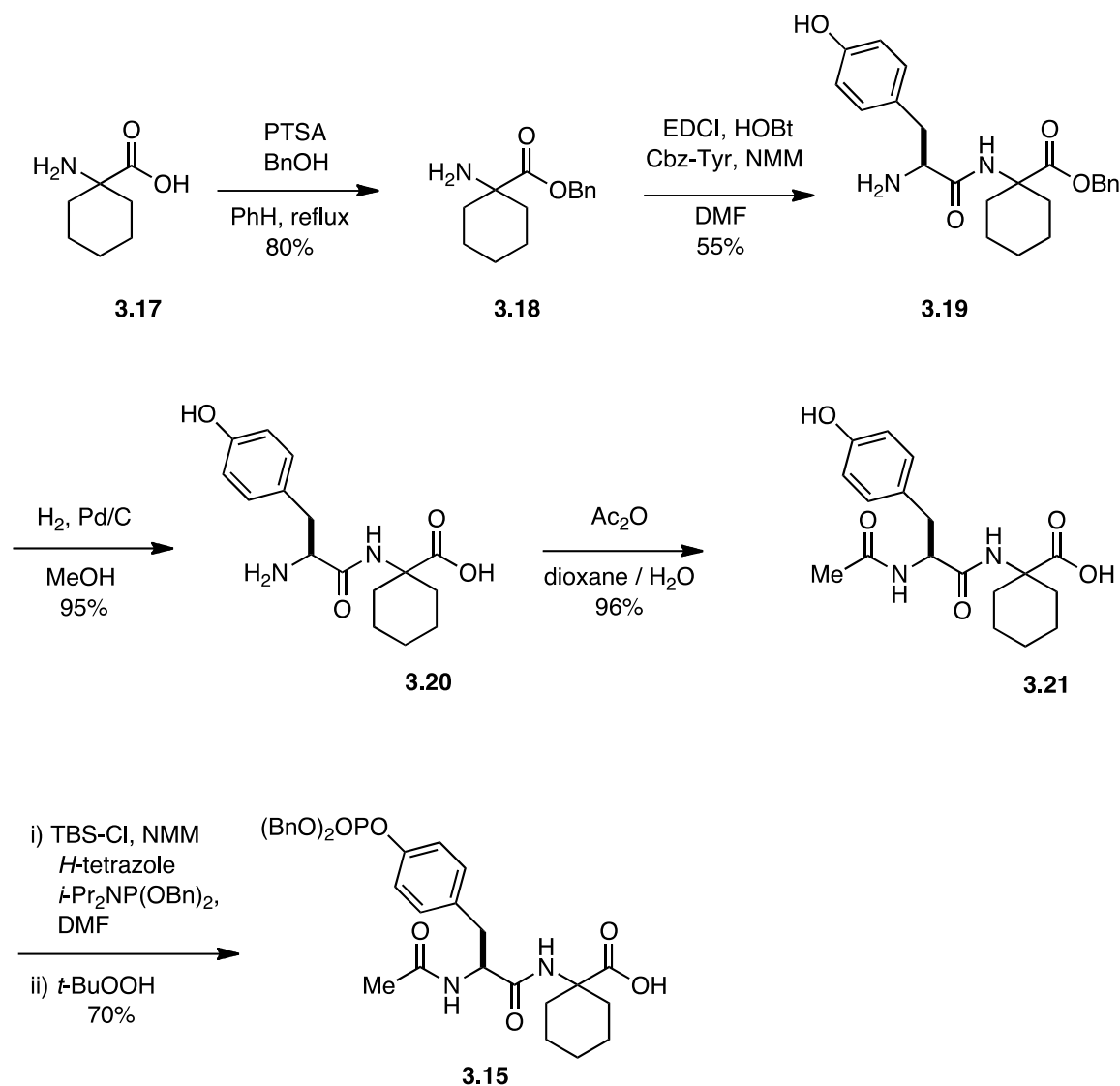
For the synthesis of ligands **3.06-3.13**, we envisioned a convergent route that features a disconnection at the pY+2 position to provide fragments A and B (Scheme 3.4.1). The advantage of this approach is that it would allow use of fragment A as a common intermediate, that may be coupled with the *N*-alkyl asparagine derivative (fragment B) to provide the final ligands in only one step. Since the residue at the pY+1 position is achiral, there is no risk of racemising the center upon a standard peptide coupling.

**scheme 3.4.1** Retrosynthesis of desired tripeptides.



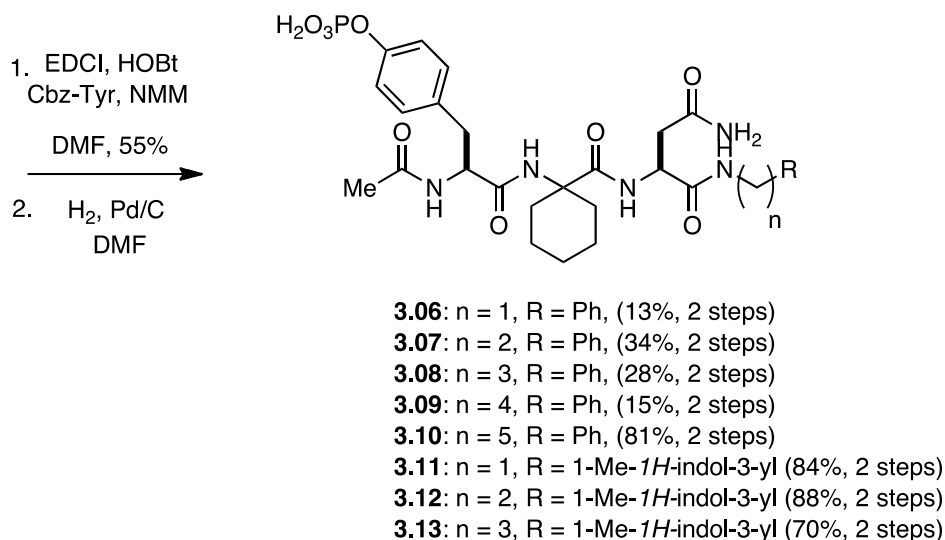
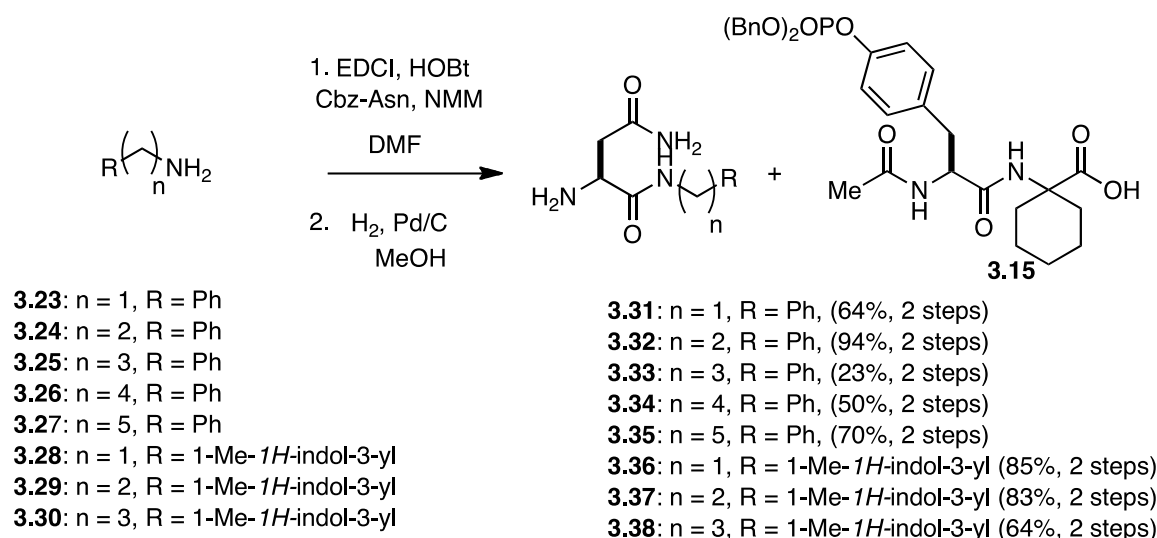
Commercially available 1-amino cyclohexyl carboxylic acid ( $\text{Ac}_6\text{c}$ ) was first converted into the known benzyl ester **3.18**. After coupling **3.15** with Cbz-Tyr, the dipeptide **3.19** thus formed was deprotected to give **3.20** in 95% yield (Scheme 3.4.2). The dipeptide was then acylated using acetic anhydride in dioxane- $\text{H}_2\text{O}$  (1:1) to provide **3.21**. Using a known procedure for the phosphorylation of tyrosine derivatives, the benzyl protected phosphorylated dipeptide **3.22** was obtained in 70% yield. Overall, this route provided large amounts of a common intermediate to provide rapid access to all of the ligands needed for this study. Indeed, the ability to stock decagram quantities of **3.22** was useful in providing rapid and convergent access to our desired compounds.

**scheme 3.4.2** Synthesis of Fragment A



The Asn derivatives that would make up **Fragment B** were synthesized via the coupling Cbz-protected Asn to amines **3.23–3.30** using carbodiimide, followed by deprotection conditions and then deprotected to give the Asn derivatives **3.27–3.34** that were coupled with dipeptide **3.19** to give **3.31–3.38** in moderate yields. Global deprotection with H<sub>2</sub> and 10% Pd/C provided the ligands **3.06–3.13** in 22–83% yields.

**scheme 3.4.3** Synthesis of desired peptides for the linker studies.

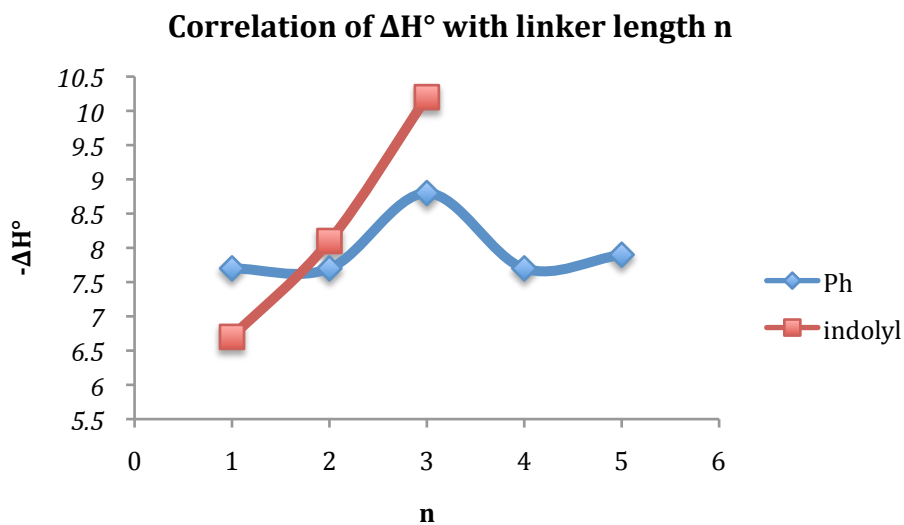


### 3.5 Isothermal Titration Calorimetry

Having ligands **3.06-3.13** now in hand, the binding energetics for the Grb2 SH2 domain were conducted. As shown in Table 3.5.1, these ITC studies indicated a clear preference in ligand binding affinity for ligands having a three carbon methylene linker *in both* the phenyl and indolyl series. Namely, there is a maximum in both  $K_a$  and  $\Delta H^\circ$  at  $n = 3$  for the phenyl substituted ligands, and for the indolyl series, there is a maximum in

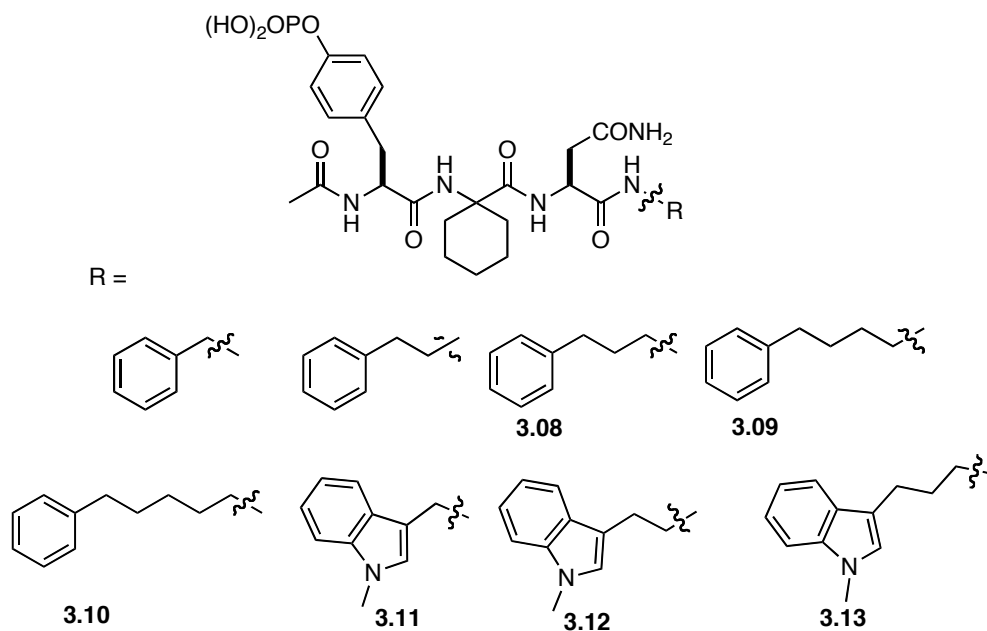
affinity for  $n = 3$ . This result supports the previous hypothesis based on molecular docking studies that a trimethylene unit provided optimal binding affinities for the pY+3 derivatives. We were also intrigued by the observation that the peak in binding affinity at  $n = 3$  (R = Ph) correlated very well with an increase in favorable  $\Delta H^\circ$  that decreased when  $n > 3$ . Thus, the peak in enthalpy of binding shown in Graph 3.5.1 provided preliminary evidence of a plausible increase in vdW interactions between the ligand and the protein.

**graph 3.5.1** Graph depicting the correlation of enthalpy with linker length  $n$ .





**Table 3.5.1** ITC binding studies with Grb2 SH2, at 298 K in HEPES buffer, pH 7.5. <sup>a</sup>ITC performed by Dr. Bo Cheng.



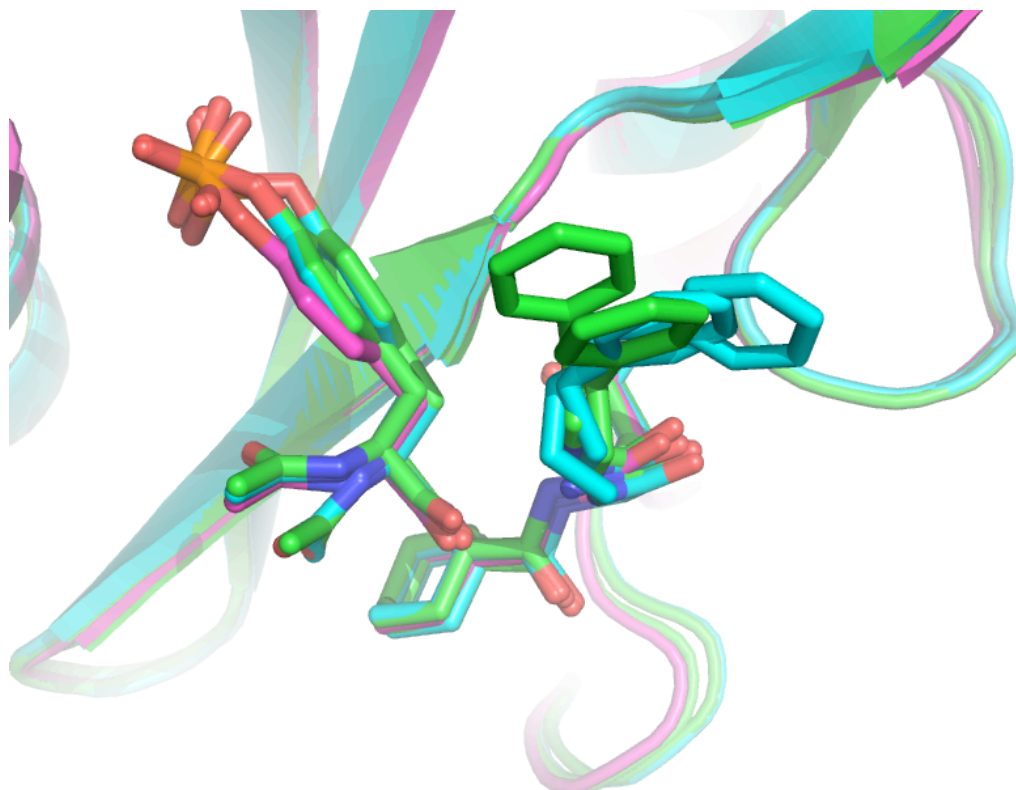
<i>Ligand</i>	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )	$-T\Delta S$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )
<b>3.06</b>	$1.5 \pm 0.2$	$-9.8 \pm 0.1$	$-7.7 \pm 0.3$	$-2.1 \pm 0.8$
<b>3.07</b>	$2.6 \pm 0.2$	$-9.8 \pm 0.2$	$-7.7 \pm 0.3$	$-2.1 \pm 0.6$
<b>3.08</b>	$6.0 \pm 0.7$	$-10.2 \pm 0.2$	$-8.8 \pm 0.3$	$-1.4 \pm 0.6$
<b>3.09</b>	$4.2 \pm 0.2$	$-9.8 \pm 0.1$	$-7.7 \pm 0.4$	$-2.0 \pm 0.5$
<b>3.10</b>	$2.8 \pm 0.1$	$-10.2 \pm 0.1$	$-7.9 \pm 0.3$	$-2.3 \pm 0.3$
<b>3.11<sup>a</sup></b>	$1.1 \pm 0.2$	$-9.6 \pm 0.1$	$-6.7 \pm 0.2$	$-2.9 \pm 0.3$
<b>3.12<sup>a</sup></b>	$2.4 \pm 0.1$	$-10.1 \pm 0.1$	$-8.1 \pm 0.2$	$-2.0 \pm 0.2$
<b>3.13<sup>a</sup></b>	$4.9 \pm 0.1$	$-10.4 \pm 0.1$	$-10.2 \pm 0.2$	$-0.2 \pm 0.3$

### 3.6 Structural Studies of the pY+3 ligands of Varying Linker Length $n$

Since one of the initial goals of the study was to examine how changes in binding affinity accompany the addition of *nonpolar surface area* to a ligand, the thermodynamic studies of varying the linker length  $n$  generated a great amount of curiosity. Accordingly, to better understand our previous ITC data in a fashion that would shed light onto the source of the  $\Delta H^\circ$  values, we obtained crystal structures of complexes of **3.07** and **3.08** with the Grb2 SH2 domain. All X-ray data was collected by Dr. John Clements. These structures were obtained to a resolution of  $<2.0$  Å, and backbone atoms belonging to the protein in the complexes of the ligands align closely, with pair wise RMSDs  $<0.2$  Å. The two structures are shown below in

Figure 3.6.1 overlaid with complex **3.04** where R = H for a comparison. Interestingly, the peptide backbone of **3.07** and **3.08** are aligned closely to one another, but the phenyl substituent at the pY+3 position of **3.07** binds in two poses, one of which positions the phenyl ring away from the protein relative to the other. The phenyl substituent of **3.08** also binds in two poses, however, in both of these the phenyl rings are localized to the non polar residues of the extended “hydrophobic patch” of the domain. This result can be rationalized such that the three-carbon linker allows for better spatial position of the phenyl substituent to make favorable vdW contacts with the binding pocket.

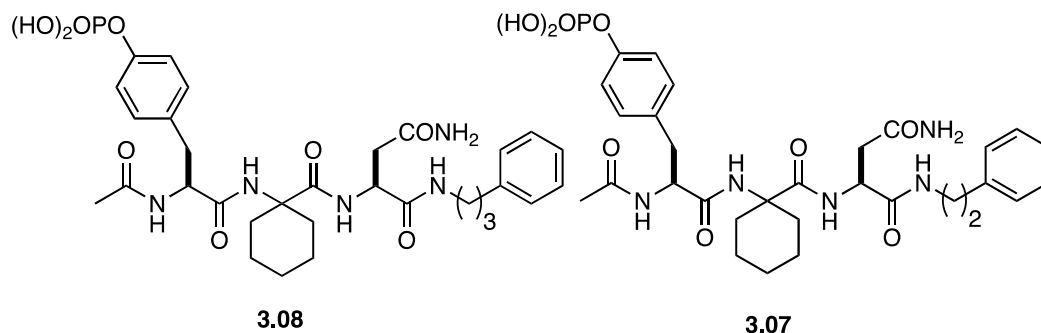
**Figure 3.6.1** Cocrystal of the Grb2 SH2 domain (surface).<sup>[a]</sup>



[a] The binding site of the Grb2 SH2 domain. Ligand **3.07** is in green and **3.08** in cyan and ligand **3.04** is in magenta (shown in sticks). Only carbon atoms are shown.

The vdW contacts made between the protein and the pY+3 substituents for both **3.07** and **3.08** were tallied ( between the distances of 3.4 and 4.2 Å<sup>clxiii</sup>) and are shown in Table 3.6.1. In one pose, **3.07** ( $n = 2$ ) makes 13 contacts with the pocket while the other pose makes only four. On the other hand, **3.08** ( $n = 3$ ) makes eight contacts with the binding pocket in both poses. These values suggest that there may be less variability in the position of the phenyl ring at the pY+3 substituent of **3.08** than for **3.07** in their respective isoenergetic poses.

**Table 3.6.1** van der Waals contacts for ligands **3.07** and **3.08**.



<i>Ligand</i>	$\Delta G$ (kcal•mol <sup>-1</sup> )	$\Delta H$ (kcal mol <sup>-1</sup> )	$T\Delta S$ (kcal•mol <sup>-1</sup> )	<i>vdW contacts made</i>
<b>3.07</b>	-9.8 ± 0.2	-7.7 ± 0.3	-2.1 ± 0.6	13,4
<b>3.08</b>	-10.2 ± 0.2	-8.8 ± 0.3	-1.4 ± 0.6	8,8

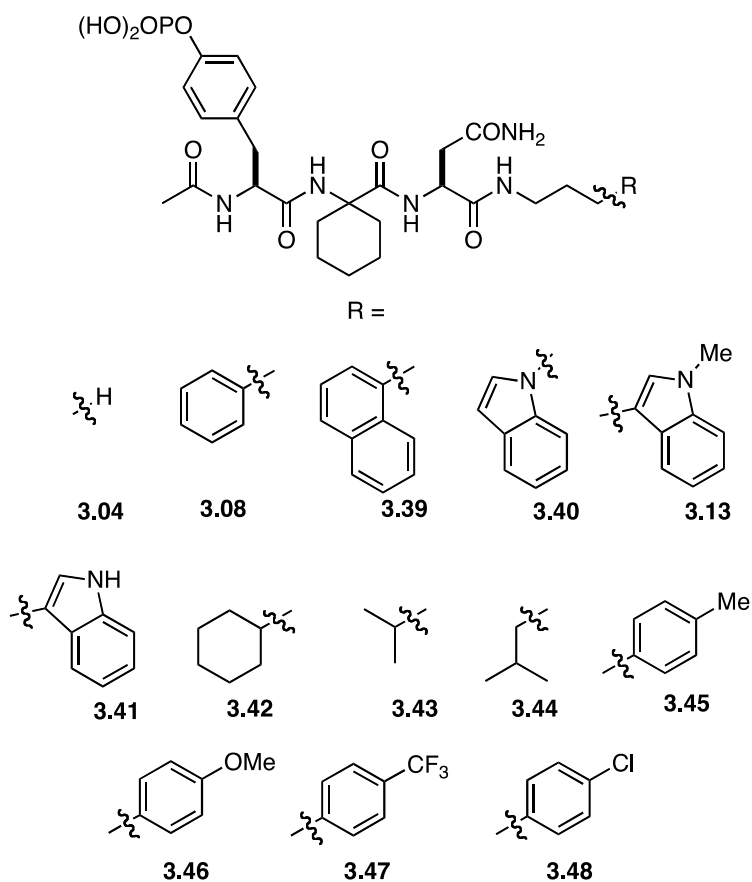
### 3.7 Summary

We examined effects upon binding affinity of varying the number of methylene groups linking the C-terminus of Grb2 SH2 domain binders with nonpolar substituents. We determined that a methylene linker of three carbon atoms was needed to provide optimal interactions between an apolar substituent R and an apolar patch of the Grb2 SH2 domain. We also obtained X-ray data suggesting that the methylene linker where  $n = 3$  provided the best spatial position of the aryl substituent to make optimal interactions with the hydrophobic patch. We would thus use a linker of  $n = 3$  in our design of additional ligands.

### 3.8 Substituent Effects at the pY+3 Derived Tripeptides.

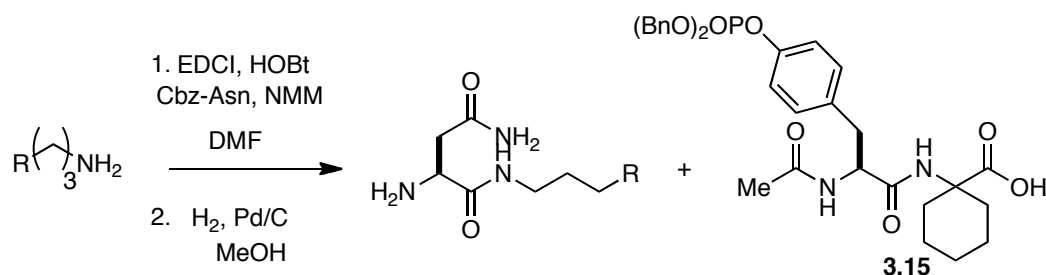
With the linker length known, we proceeded to study the thermodynamic and structural effects of varying that nature of nonpolar substituents at the pY+3 position of Grb2 SH2 binding ligands. The synthesis of all ligands studied are presented in one section even though each of the ligands shown in Figure 3.8.1 was selected to ask specific questions that will be discussed in due course. In general, these ligands contain substituents that vary in their aromatic (where R = phenyl, naphthyl, indolyl), and aliphatic (where R = cyclohexyl, isopropyl and, isobutyl) substituents. The substitution patterns on the phenyl ring were also varied to compare the effects associated with electron withdrawing *and* donating functionality. Furthermore, known ligand **3.04** (where R = H) was used as a benchmark to judge the effects of adding apolar surface area.

**Figure 3.8.1** pY+3 substituents where R is varied on a three carbon linker.



Ligands in Figure 3.3.3 were synthesized in the same manner as the previous section. In particular, the asparagine derivatives were synthesized in the same manner as in Scheme 3.2.3 and coupled with dipeptide **3.15** in moderate yields (scheme 3.3.4), and then deprotected with  $H_2$  and 10% Pd/C in MeOH to provide the debenzylated ligands **3.39-3.48**.

**Scheme 3.8.1** Peptide Synthesis.<sup>a</sup> Dr. Bo Cheng.<sup>b</sup> Dr. J. Tian.



**3.50:** R = Naphthyl

**3.51<sup>a</sup>:** R = 1*H*-indol-1-yl

**3.52<sup>a</sup>:** R = 1-*H*-indol-3-yl

**3.53:** R = Cyclohexyl

**3.54:** R = *isobutyl*

**3.55:** R = *isoPropyl*

**3.56<sup>b</sup>:** R = 4-Cl-phenyl

**3.57<sup>b</sup>:** R = 4-Me-phenyl

**3.58:** R = 4-CF<sub>3</sub>-Phenyl

**3.59:** R = 4-Ome-Phenyl

**3.60:** R = Naphthyl, (80%, 2 steps)

**3.61<sup>a</sup>:** R = 1*H*-indol-1-yl, (54%, 2 steps)

**3.62<sup>a</sup>:** R = 1-*H*-indol-3-yl (70%, 2 steps)

**3.63:** R = Cyclohexyl (72%, 2 steps)

**3.64:** R = *isobutyl* (67%, 2 steps)

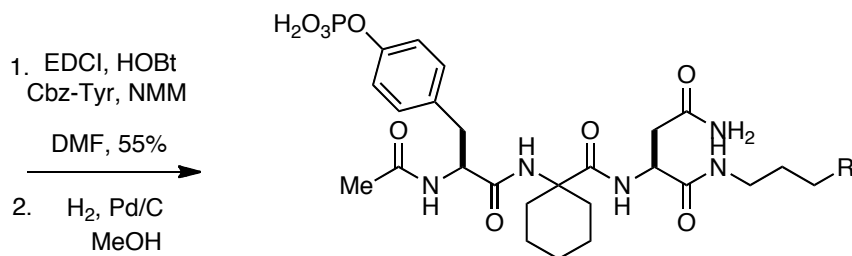
**3.65:** R = *isoPropyl* (99%)

**3.66<sup>b</sup>:** R = 4-Cl-phenyl (70%, 2 steps)

**3.67<sup>b</sup>:** R = 4-Me-phenyl (39%, 2 steps)

**3.68:** R = 4-CF<sub>3</sub>-Phenyl (65%, 2 steps)

**3.69:** R = 4-OMe-Phenyl (99%, 2 steps)



**3.39:** R = Naphthyl, (75%, 2 steps)

**3.40<sup>a</sup>:** R = 1*H*-indol-1-yl, (78%, 2 steps)

**3.41<sup>a</sup>:** R = 1-*H*-indol-3-yl (55%, 2 steps)

**3.42:** R = Cyclohexyl (75%, 2 steps)

**3.43:** R = *isoPropyl* (22%, 2 steps)

**3.44:** R = *isoButyl* (87%, 2 steps)

**3.45<sup>b</sup>:** R = 4-Me-phenyl (30%, 2 steps)

**3.46<sup>b</sup>:** R = 4-Cl-phenyl (32%, 2 steps)

**3.47:** R = 4-CF<sub>3</sub>-Phenyl (94%, 2 steps)

**3.48:** R = 4-OMe-Phenyl (20%, 2 steps)

### 3.9 Thermodynamic studies

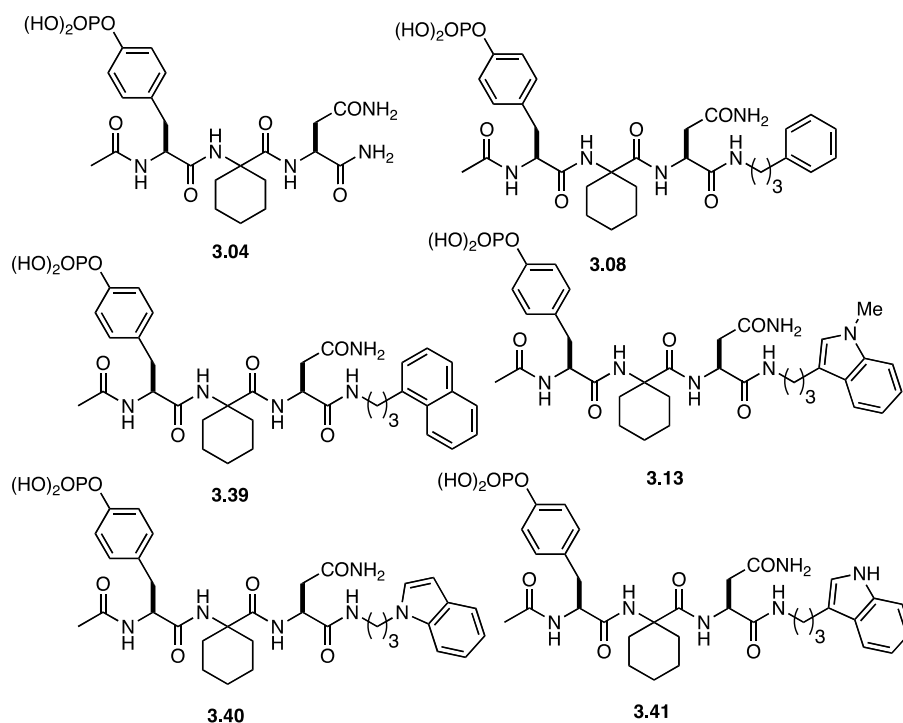
Binding energetics for ligands **3.39-3.41** with the Grb2 SH2 domain were determined by ITC (**Table 3.9.1**). When *R* is phenyl, the ligand binds with more favorable *enthalpy* and *entropy* relative to **3.04**, our benchmark ligand for comparison.



These findings indicate that the addition of apolar surface area leads to an enthalpically driven affinity enhancement, which is counterintuitive to the hydrophobic effect, which would suggest that the increase of apolar surface area would lead to potency increases that were driven by a favorable entropy. There is also a more favorable binding entropy for this event as well. The naphthyl and indolyl derived ligands **3.28** and **3.13**, respectively, bound with a more favorable change in enthalpy relative to **3.08**, but there were no significant differences in entropy changes observed. Thus, the increase in apolar surface area from R = Ph (**3.08**) to R = Np (**3.28**) led to favorable changes in binding enthalpy that resulted in an overall increase in binding affinity. Furthermore, relative to our benchmark **3.04**, we also observe with **3.28** that the affinity increases due to a more favorable binding enthalpy, but to a greater extent, which correlates well with the addition of the larger surface of **3.28** relative to **3.08**.

The indolyl ligands **3.29** and **3.30** had a slightly *less* favorable enthalpy of binding relative to **3.08**. We made numerous attempts to obtain a crystal structure of this complex, but all attempts have thus far failed.

**Table 3.9.1** ITC binding studies with Grb2 SH2, at 298 K in HEPES buffer, pH 7.5.



Ligand	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta H^\circ$ ( $\text{kcal mol}^{-1}$ )	$-T\Delta S^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )
<b>3.04</b>	$0.7 \pm 12$	$-9.3 \pm 0.1$	$-8.5 \pm 0.4$	$-0.8 \pm 0.4$
<b>3.08</b>	$6.0 \pm 0.7$	$-10.2 \pm 0.2$	$-8.8 \pm 0.3$	$-1.4 \pm 0.6$
<b>3.39</b>	$6.0 \pm 0.3$	$-10.6 \pm 0.1$	$-9.6 \pm 0.4$	$-1.2 \pm 0.5$
<b>3.13<sup>a</sup></b>	$4.9 \pm 0.1$	$-10.4 \pm 0.1$	$-10.2 \pm 0.2$	$-0.2 \pm 0.3$
<b>3.40<sup>a</sup></b>	$2.7 \pm 0.2$	$-10.2 \pm 0.3$	$-9.1 \pm 0.8$	$-1.1 \pm 0.8$
<b>3.41<sup>a</sup></b>	$3.5 \pm 0.2$	$-10.3 \pm 0.1$	$-9.1 \pm 0.3$	$-1.2 \pm 0.6$

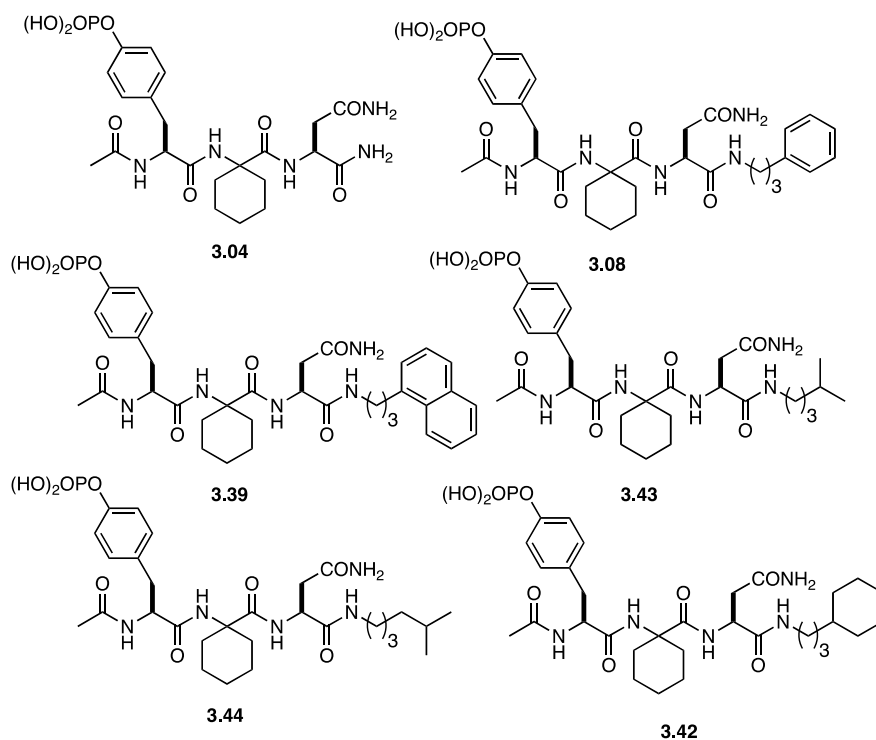
<sup>a</sup>Performed by Dr. Bo Cheng. TC experiments were conducted at 25 °C in HEPES buffer at pH  $7.45 \pm 0.05$ . Errors based on 5% error in ligand concentration plus standard deviations.

At this point in our studies, it was clear that we were observing *enthalpically* driven hydrophobic effects upon the addition of aromatic surface area to the test ligand **3.04**, but an explanation as to where the enthalpy originated was lacking. It was conceivable that the enhanced change in binding entropy arose from favorable dispersive interactions with C-terminal substituents, but there was little support for this hypothesis at that time. Moreover, we were seeing more favorable gains in both entropy and enthalpy relative to benchmark **3.04** for the binding event. These observations lead us to believe that there was not just one contributing factor at play, and thus we modified our studies to include structural and thermodynamic analysis of the alkyl ligands **3.42–3.44**. Since cyclohexane has been postulated to have a more entropically favorable dehydration in water than benzene,<sup>clxiv</sup> we reasoned that studying R = cyclohexane for R = benzene might lead to added hydrophobic effects that may help us better discern the effects of added apolar surface area.

Accordingly, with the alkyl ligands **3.42–3.44** series in hand, the ITC data was collected for their binding to the Grb2 SH2 domain (Table 3.9.2). For these ligands, it was observed that they bind with more favorable  $\Delta\Delta S^\circ$  compared to **3.08**, but this is wiped out by a less favorable binding  $\Delta\Delta H^\circ$  such that **3.42–3.44** so that  $\Delta\Delta G^\circ$  relative to **3.08** is unfavorable. The addition of apolar surface area still led to binding affinity increases relative to **3.04**. However, it was remarkable that whereas the aromatic ligands bound with more favorable enthalpy relative **3.04**, the alkyl ligands bound with more favorable entropy. This indicated that the difference in the dehydration of the alkyl verse aromatic pY+3 substituents may have an effect on the binding enthalpy and entropy, but further experiments were needed to explore that hypothesis. Once such experiment would

be to compare the temperature dependent thermodynamics, or  $\Delta Cp^\circ$ , of the alkyl and aromatic ligands. In biological systems,  $\Delta Cp$  is frequently used as a barometer of hydrophobic effects.<sup>clxv</sup> A negative value of  $\Delta Cp$  is associated with the burial of nonpolar surface area and a positive number with the burial of polar surface area. Thus, this analysis could provide additional data to discern if the alkyl substituents were burying more apolar surface area compared to **3.08**. The analyses of this experiment is discussed in the following section.

**Table 3.9.2** ITC binding studies with Grb2 SH2.



<i>Ligand</i>	$K_a$ (x 10 <sup>7</sup> M <sup>-1</sup> )	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )
<b>3.04</b>	0.7 ± 12	-9.3 ± 0.1	-8.5 ± 0.4	-0.8 ± 0.4
<b>3.08</b>	6.0 ± 0.7	-10.2 ± 0.2	-8.8 ± 0.3	-1.4 ± 0.6
<b>3.39</b>	6.0 ± 0.3	-10.6 ± 0.1	-9.6 ± 0.4	-1.2 ± 0.5
<b>3.43</b>	1.7 ± 0.2	-9.8 ± 0.1	-7.8 ± 0.4	-2.1 ± 0.3
<b>3.44</b>	1.7 ± 0.2	-9.8 ± 0.2	-7.8 ± 0.2	-2.1 ± 0.2
<b>3.42</b>	1.9 ± 0.2	-9.8 ± 0.1	-8.0 ± 0.4	-1.9 ± 0.2

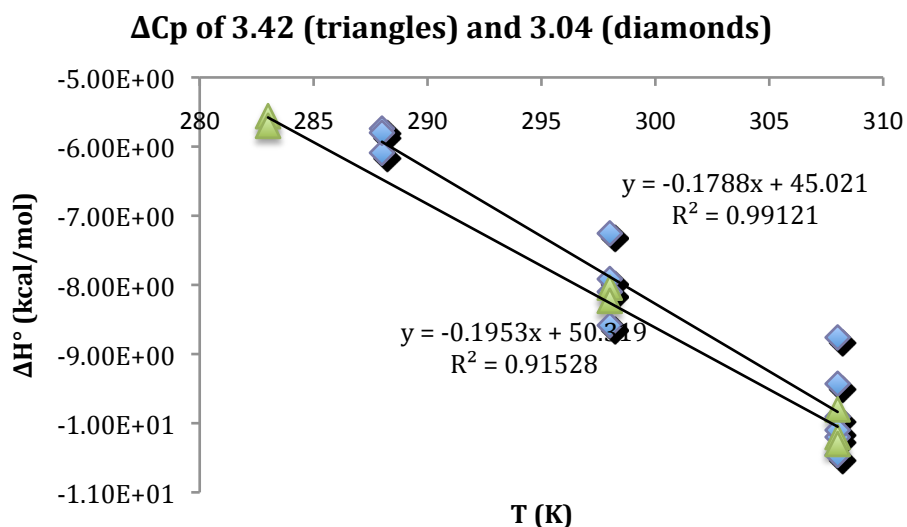
<sup>a</sup>

ITC experiments were conducted at 25 °C in HEPES buffer at pH 7.45 ± 0.05. Errors based on 5% error in ligand concentration plus standard deviations.

### 3.10 Temperature Dependent Studies

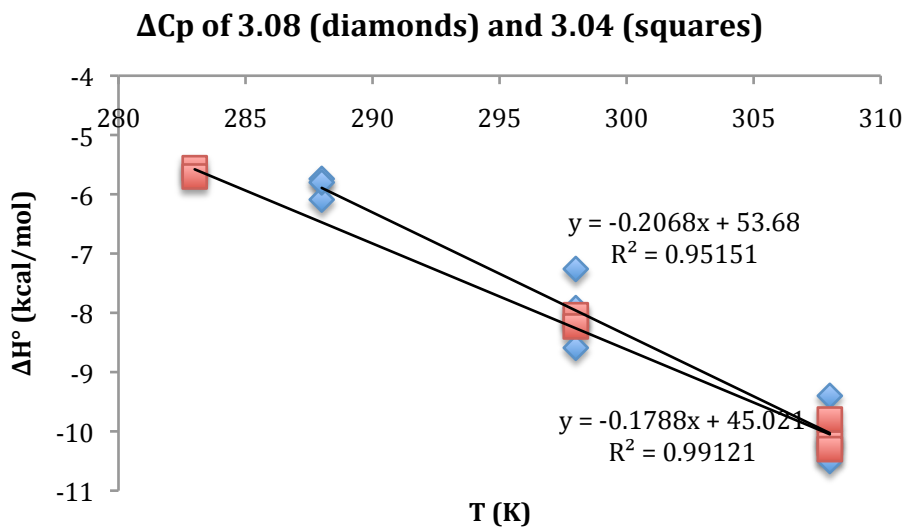
The  $\Delta C_p$  values of the aromatic ligands **3.08**, **3.28**, and **3.13** are shown in figure 3.10.1-4. The  $\Delta C_p$  for **3.04** is also shown on each respective  $\Delta C_p$  trace for comparison.

**Figure 3.10.1**  $\Delta H^\circ$  of binding to the Grb2 SH2 domain as a function of temperature.



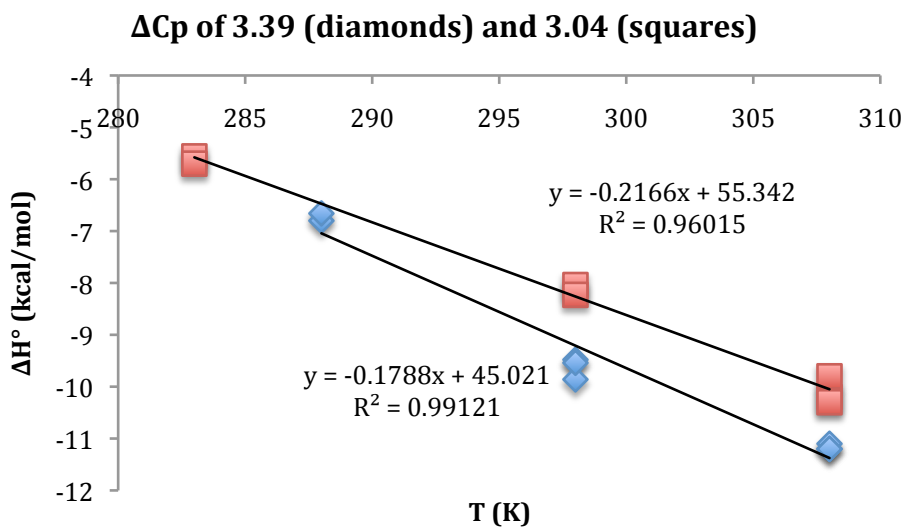
Measurements were performed in 50 mM HEPES buffer containing 150 mM NaCl at pH 7.5.  $\Delta C_p$  values for the binding of each ligand were obtained from the slope of the plots and the error in  $\Delta C_p$  was taken as the standard error in the slope. All runs for 3.04 were done by Dr. James Myslinski

**Figure 3.10.2**  $\Delta H^\circ$  of binding to the Grb2 SH2 domain as a function of temperature for **3.08**



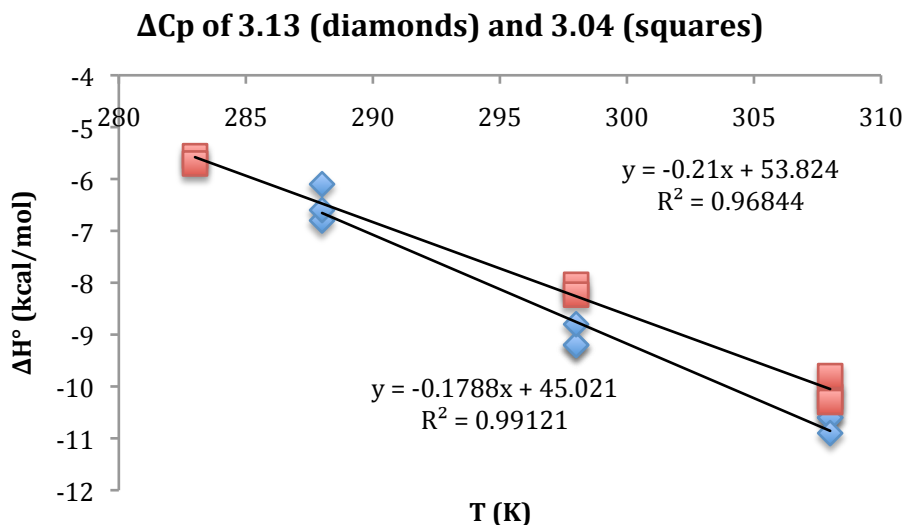
Measurements were performed in 50 mM HEPES buffer containing 150 mM NaCl at pH 7.5.  $\Delta C_p$  values for the binding of each ligand were obtained from the slope of the plots and the error in  $\Delta C_p$  was taken as the standard error in the slope. All runs for 3.04 were done by Dr. James Myslinski.

**Figure 3.10.3**  $\Delta H^\circ$  of binding to the Grb2 SH2 domain as a function of temperature for **3.28**



Measurements were performed in 50 mM HEPES buffer containing 150 mM NaCl at pH 7.5.  $\Delta C_p$  values for the binding of each ligand were obtained from the slope of the plots and the error in  $\Delta C_p$  was taken as the standard error in the slope. All runs for 3.04 were done by Dr. James Myslinski.

**Figure 3.10.4**  $\Delta H^\circ$  of binding to the Grb2 SH2 domain as a function of temperature for **3.13**



Measurements were performed in 50 mM HEPES buffer containing 150 mM NaCl at pH 7.5 by Dr. Bo Cheng.  $\Delta C_p$  values for the binding of each ligand were obtained from the slope of the plots and the error in  $\Delta C_p$  was taken as the standard error in the slope. All runs for **3.04** were done by Dr. James Myslinski.

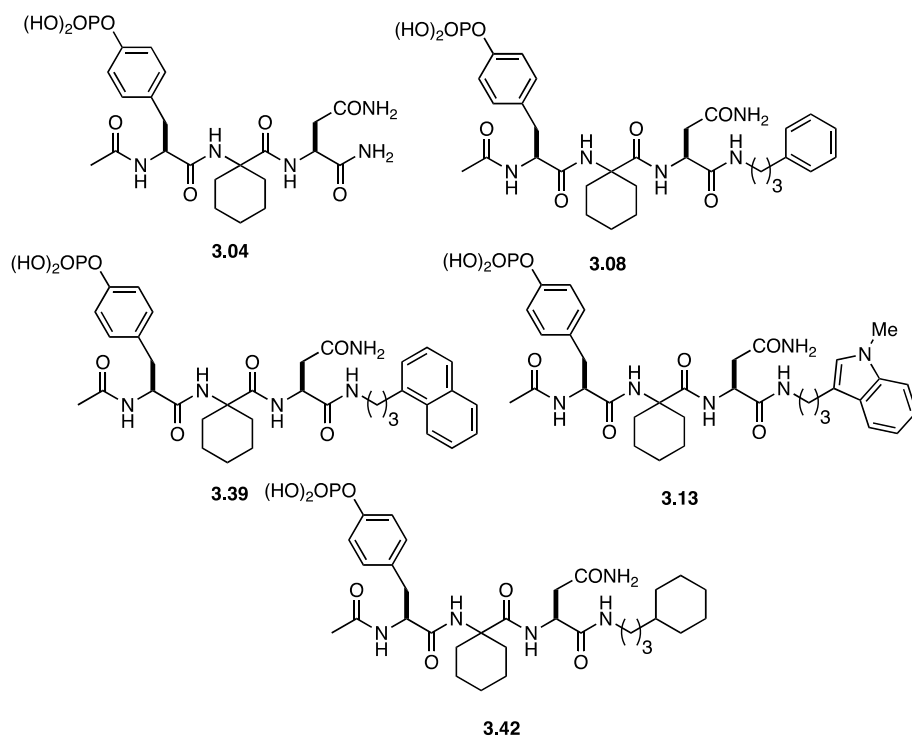
These experiments show that when R is aromatic, the  $\Delta C_p$  values for **3.08**, **3.39**, and **3.13** become more negative upon addition of the aromatic substituent, which correlates well with the increase in binding affinity as extra nonpolar surface is added to **3.04**, Table 3.10.1. These are in agreement with what one may expect from targeting classical hydrophobic effects to enhance to potency of a drug lead; however, the burial of nonpolar surface area correlates with a favorable change in  $\Delta H^\circ$  and *not*  $\Delta S^\circ$  (Table 3.10.1), which indicates a hydrophobic effect driven by enthalpy. Since hydrophobic effects are supposed to increase binding affinity due to a more favorable entropic component, this is counter to what is expected.

Because the  $\Delta C_p$  data of **3.08**, **3.28**, and **3.13** confirmed that the addition of nonpolar surface area correlated well with more favorable changes in binding enthalpy of the binding event (Table 3.10.1), we then collected  $\Delta C_p$  data for a respective ligand **3.42**



where R is aliphatic to determine if it trended similarly. Thus, the  $\Delta C_p$  for **3.42** was determined and is also shown with **3.04** included as a comparison. We observed that **3.42**, similar to ligands **3.08**, **3.28**, and **3.13** has a  $\Delta C_p$  that is more negative than the benchmark **3.04**. The value for the  $\Delta C_p$  of **3.42** suggests that we are observing a hydrophobic effect upon the addition of apolar surface area to **3.04**, which is expected given what is known for hydrophobic effects.

**Table 3.10.1 All acquired values for  $\Delta Cp^\circ$ .**



<i>Ligand</i>	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G^\circ$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta Cp^\circ$
<b>3.04</b>	$0.7 \pm 12$	$-9.3 \pm 0.1$	$-179 \pm 10$
<b>3.08</b>	$6.0 \pm 0.7$	$-10.2 \pm 0.2$	$-209 \pm 12$
<b>3.39</b>	$6.0 \pm 0.3$	$-10.6 \pm 0.1$	$-217 \pm 8$
<b>3.13</b>	$4.9 \pm 0.1$	$-10.4 \pm 0.1$	$-210 \pm 10$
<b>3.42</b>	$1.9 \pm 0.2$	$-9.8 \pm 0.1$	$-196 \pm 8$

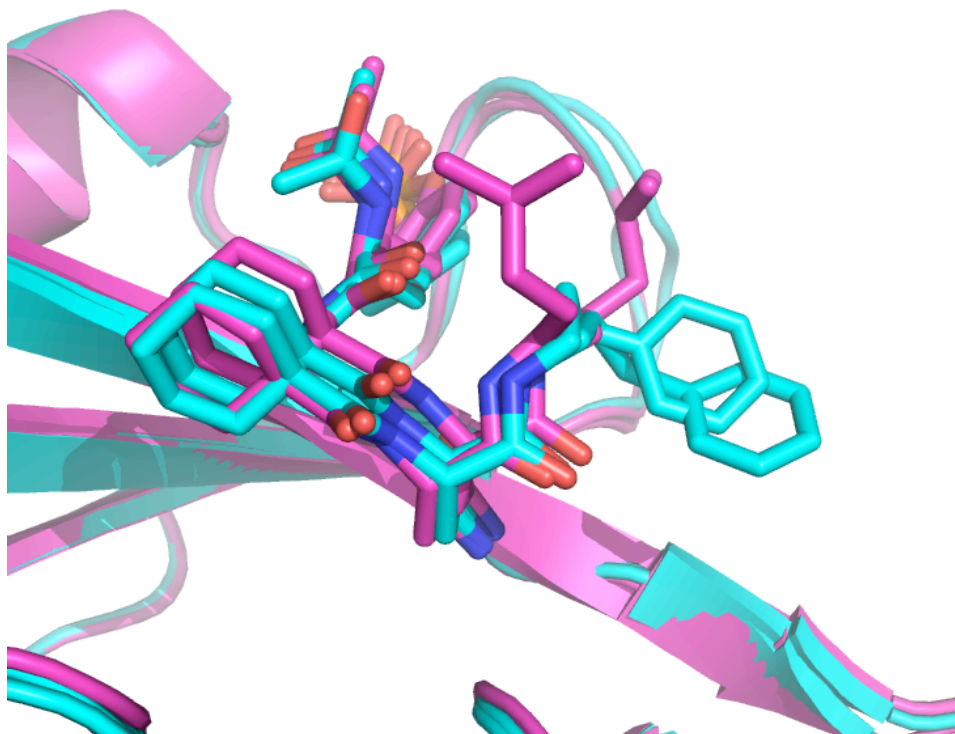
### 3.11 Structural Studies.

Unfortunately, we were not able to obtain an X-ray structure for **3.42** (R = Cy), which provides the most appropriate comparison to **3.08** because they have the same number of carbon atoms. However, we were able to obtain a crystal structures of **3.43** (R = *i*Pr) complexed with Grb2 SH2 domain, and it is shown in Figure 3.11.1 overlaid with **3.08** as a reference. The structure was obtained to a resolution of <2.0 Å and backbone atoms belonging to the protein in the complexes of **3.43** and **3.04** align closely, with pairwise RMSDs <0.2 Å. Analysis of the X-ray data shows that the orientation of the *i*Pr substituent at the pY+3 position **3.43** varies drastically, and the *i*Pr group in two different poses. In both of these, the *i*Pr group is directed away from the hydrophobic patch of Grb2. An analysis of the van der Waals contacts made at the pY + 3 position (using the distance criteria of 3.4-4.2 Å for nonbonded atoms) is shown in Table 3.11.1 in addition to the contacts made by **3.08** as a reference. As previously discussed in Section 3.6, the phenyl substituent in **3.08** binds in an orientation so that it makes eight favorable contacts (in each pose) with the domain. However, **3.43** binds with fewer vdW contacts with three contact in one pose and two in the other pose.

There was more variation in the orientation of the *i*Pr group in **3.43** than the Ph group in **3.08**, and we noticed this trend correlates well with the differences in binding entropy observed for the alkyl ligands. Furthermore, there is also an apparent structural correlation between the eight vdW contacts made in **3.08** and its binding enthalpy relative to our benchmark. Thus, the data suggests that the flat, more rigid aromatic substituent may better suited to make more favorable

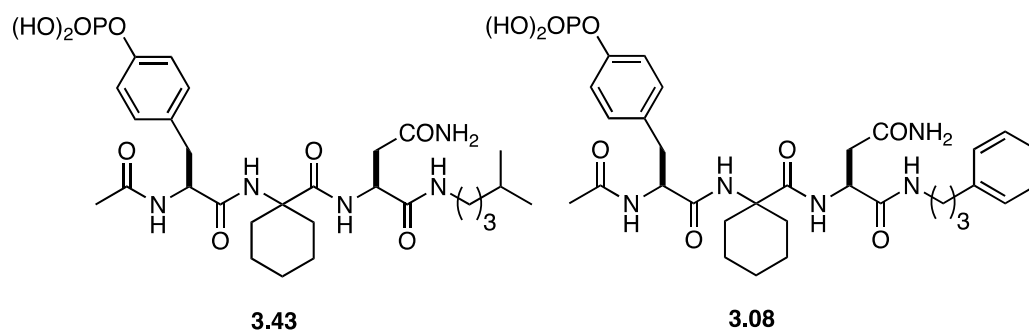
vdW contacts than the alkyl substituents, at least one of which (**3.43**) has been shown through structural analysis to have varying position in the binding complex. This is in agreement with our observation that the alkyl ligands bind with an unfavorable  $\Delta\Delta H^\circ$  between 1.0–1.8 kcal•mol<sup>-1</sup> relative to the aromatic series. Thus, we reasoned that the more rigid, aromatic substituents of **3.08**, **3.39**, and **3.13** interact with the domain with more dominating and favorable binding enthalpies relative to the alkyl ligands (**3.42**, **3.43**, **3.44**) due to the ability of the aromatic substituent to make more favorable dispersive interactions in the binding pocket.

**Figure 3.11.1** Cocrystal of the Grb2 SH2 domain (surface).<sup>[a]</sup>



[a] The binding site of the Grb2 SH2 domain. Ligand **3.43** is in green and **3.08** is in cyan. Only carbon atoms are shown. Credit to Dr. John Clements is given for obtaining the crystal.

**Table 3.11.1** Tabulated vdW contacts.



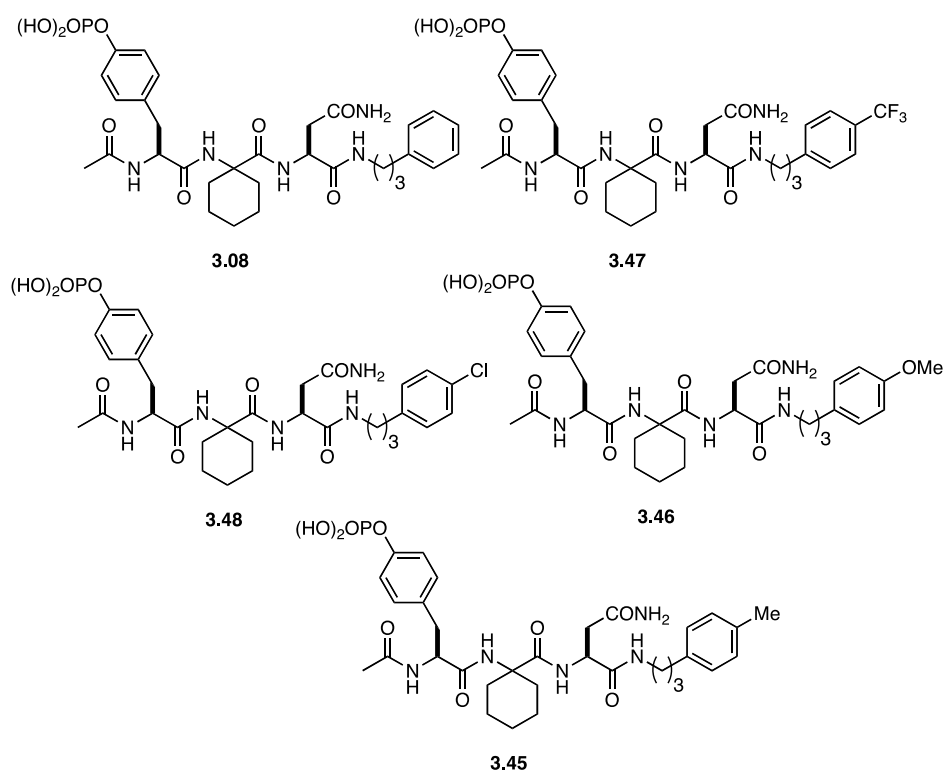
Ligand	$\Delta G^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H^\circ$ (kcal mol <sup>-1</sup> )	$-T\Delta S^\circ$ (kcal•mol <sup>-1</sup> )	vdW contacts made
<b>3.08</b>	$-10.2 \pm 0.2$	$-8.8 \pm 0.3$	$-1.4 \pm 0.6$	8, 8
<b>3.43</b>	$-9.8 \pm 0.1$	$-7.8 \pm 0.4$	$-2.1 \pm 0.3$	3, 2

### 3.12 Phenyl Substituent effects

We were also interested in probing the effects of altering the electrostatic nature of the phenyl group of **3.08**, and towards this end **3.45** – **3.48** were designed. As our inspiration for selecting these ligands, we chose phenyl substituents from the Topliss operational schemes<sup>clxvi</sup> whereby a small set of phenyl-substituted analogs are synthesized and their measured binding affinities guide the synthesis to a compound with optimized potency. These choices in synthesis are guided by considerations of hydrophobicity,<sup>clxvii</sup> steric effects,<sup>clxviii</sup> and electrostatics<sup>cx</sup> of different substituents. We hoped that by applying these schemes to **3.08** that we might obtain trends in the substituent effects that might assist in our interpretation of the binding thermodynamics.

These ligands **3.34–3.37** were synthesized (Section **3.8**) and ITC data in the presence of the Grb2 SH2 domain were collected and is shown in Table 3.12.1. Examination of these results reveal that modifying the substitution on the phenyl ring has little effect on the ligand binding affinity. However, out of the group of “Topliss” ligands that were tested, the 4-CF<sub>3</sub> derived **3.34** had the most favorable entropy of binding. This finding is in agreement with previous reports<sup>clxix</sup> suggesting that increasing the van der Waal radii of phenyl substituents can lead to a more favorable *entropy* of binding relative to the unsubstituted or starting phenyl-derived ligand. This effect has been attributed to a more favorable entropy of ligand desolvation due to the larger substituent on the phenyl ring. Although this does not help us better discern the binding interactions of this system, it is at least a noteworthy observation.

**Table 3.12.1** ITC binding studies with Grb2 SH2, at 298 K in PBS buffer, pH 7.5.  
<sup>a</sup>Synthesized and tested by Dr. J. Tian.



Ligand	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G$ (kcal $\cdot$ mol <sup>-1</sup> )	$\Delta H$ (kcal mol <sup>-1</sup> )	$-T\Delta S$ (kcal $\cdot$ mol <sup>-1</sup> )
<b>3.08</b>	$6.0 \pm 0.7$	$-10.2 \pm 0.2$	$-8.8 \pm 0.3$	$-1.4 \pm 0.6$
<b>3.46</b>	$4.7 \pm 0.5$	$-10.3 \pm 0.1$	$-9.2 \pm 0.6$	$-1.2 \pm 0.7$
<b>3.48<sup>a</sup></b>	$5.1 \pm 0.4$	$-10.5 \pm 0.1$	$-9.5 \pm 0.4$	$-1.0 \pm 0.6$
<b>3.47</b>	$3.0 \pm 0.7$	$-10.1 \pm 0.1$	$-8.4 \pm 0.2$	$-1.7 \pm 0.4$
<b>3.45<sup>a</sup></b>	$4.7 \pm 0.4$	$-10.3 \pm 0.1$	$-9.0 \pm 0.4$	$-1.3 \pm 0.4$

ITC experiments were conducted at 25 °C in HEPES buffer at pH  $7.45 \pm 0.05$ . Errors based on 5% error in ligand concentration plus standard deviations. <sup>a</sup>Performed by Dr. J. Tian.

### 3.13 Exploring Leveling Effects

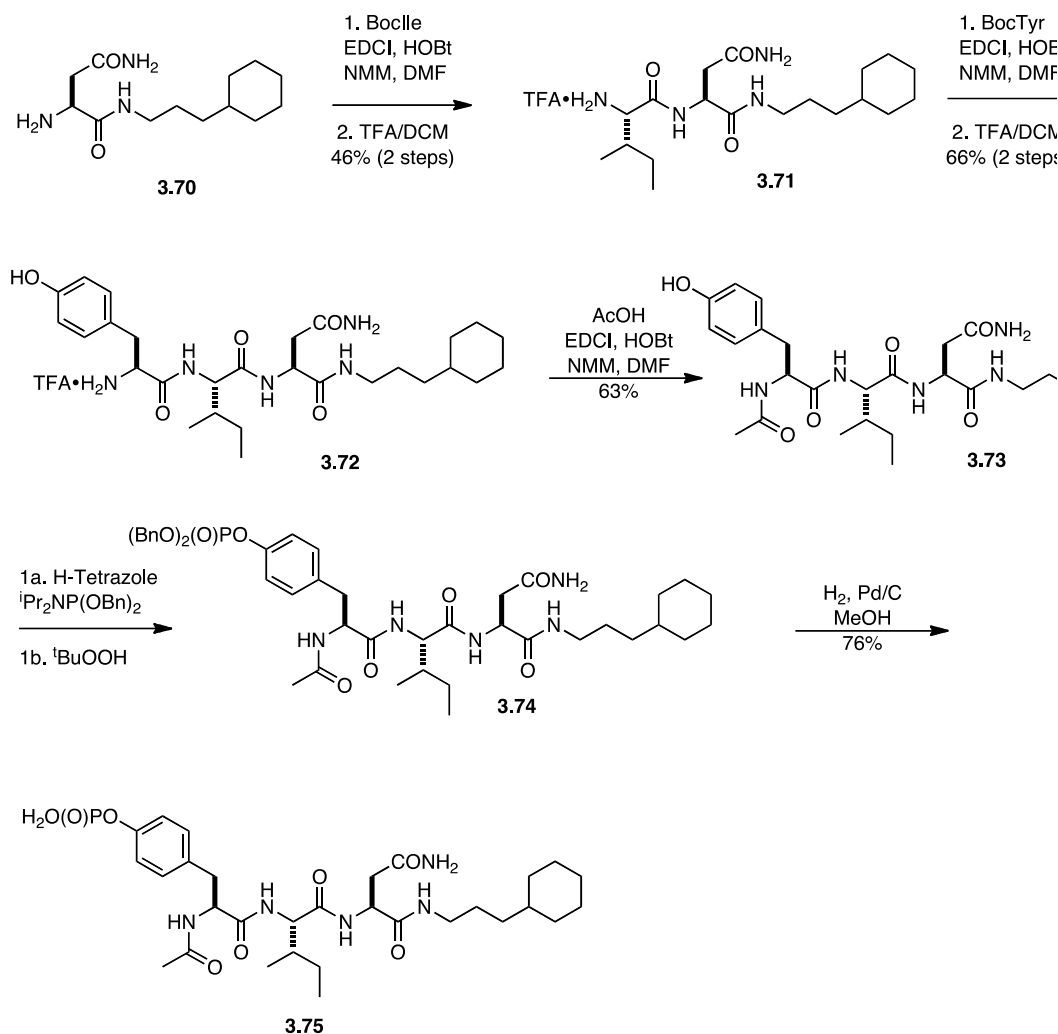
What remained puzzling to us after our phenyl substitution studies is that we were still observing enthalpy-entropy compensation rather than a significant



change in binding affinity. We wondered if this was because we had reached an affinity plateau for this protein-ligand system such that no further affinity gains could be achieved.<sup>clxx</sup> In an effort to circumvent this problem, we decided to change the pY+1 residue of our ligand framework from Ac6c to Ile, which would give us **3.76** (Table 3.13.1) as a less potent starting benchmark. Thus, we postulated that since this would give us a lower affinity starting point for optimization that we may not hit the supposed affinity plateau that we encountered in our studies. Therefore, we decided to synthesize and test a derivative of **3.76** that was modified with a cyclohexyl substituent to provide **3.75** (table 3.13.1).

Accordingly, **3.40** was coupled with NHBocIle under carbodiimide coupling conditions to give an intermediate Boc-protected dipeptide that was deprotected with TFA to give **3.71**. The following TFA salt was similarly coupled to BocTyr and deprotected to give **3.72**. The tripeptide was then acylated and phosphorylated to give **3.74**, which was then easily debenzylated under hydrogenolysis conditions to provide the desired ligand **3.75** (Scheme 3.13.1)

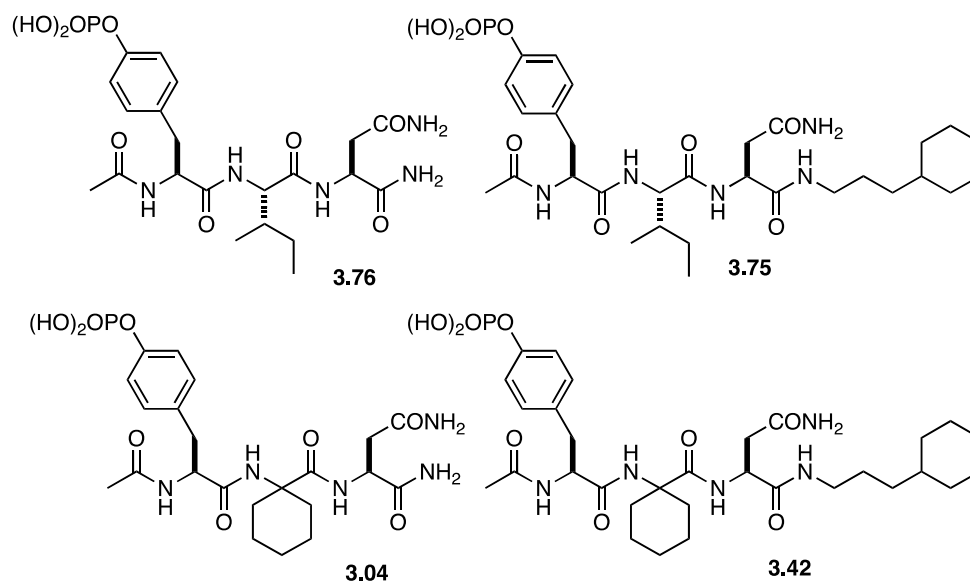
**scheme 3.13.1** pY+3 leveling effect Ligand Synthesis.



With this ligand in hand, we obtained ITC data for binding with Grb2-SH2 and they were compared to **3.76** which had already been tested by colleague Dr. James Myslinski (table 3.13.1). Interestingly, we found that **3.38** did not have a significant affinity increase compared to **3.37**, which is a puzzling result. We initially decided to make and synthesize more derivatives, but it was shown by Dr. James Myslinski on somewhat related work at the pY-1 position that showed that the substituent effects were additive and thus we did not proceed with making the other analogues.

clxxi

**Table 3.13.1** ITC analysis with Grb2 SH2 domain in HEPES at pH 7.5. Ligands made and tested by Dr. James Myslinski.



<i>Ligand</i>	$K_a$ ( $\times 10^7 \text{ M}^{-1}$ )	$\Delta G$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )	$T\Delta S$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )
<b>3.76<sup>a</sup></b>	$0.1 \pm 0.04$	$-8.2 \pm 0.02$	$-7.4 \pm 0.2$	$0.8 \pm 120.2$
<b>3.75</b>	$0.2 \pm 0.04$	$-8.3 \pm 0.1$	$-6.5 \pm 0.4$	$1.8 \pm 0.3$
<b>3.04<sup>a</sup></b>	$0.7 \pm 12$	$-9.3 \pm 0.1$	$-8.5 \pm 0.4$	$-0.8 \pm 0.4$
<b>3.42</b>	$1.9 \pm 0.2$	$-9.8 \pm 0.1$	$-8.0 \pm 0.4$	$-1.9 \pm 0.2$

### 3.14 Conclusions

We have shown that the added substituents probed in this study provided binding increases to our Grb2 SH2 ligands, but the thermodynamic driving force ( $\Delta H^\circ$  or  $\Delta S^\circ$ ) of binding affinities depended greatly upon the topical surface of the apolar substituent. Through a combination of X-ray and ITC studies it was shown that ligands containing aromatic and planar functional groups bound with a higher  $\Delta H^\circ$  than the more flexible alkyl ligands, and that this effect is consistent with more direct vdW contacts made in the pocket, based off of our limited structural analysis. We also observed that when the phenyl ring in **3.08** was replaced by more nonpolar surface area such as a naphthyl (**3.39**) or indolyl (**3.13**) substituent, that further affinity increases were driven by more favorable changes in enthalpy. Interestingly, this increase in binding affinity was also accompanied by a synergistic benefit in entropy that may be attributed to the release of ordered water molecules, although no crystallographic data was obtained to support or refute this hypothesis.

However, when we applied a similar tactic to our benchmark **2.04** by substituting with less planar, alkyl substituents, affinity boosts were observed but in this case they were driven by entropy, and they were not as large in magnitude as with the aromatic ligands. Based upon our X-ray structural analysis, the binding entropy correlated with the alkyl substituents having more flexibility in the binding pocket as well as having fewer contacts made in the binding pocket. The alkyl ligand examined was in two drastically different poses that we observed

by X-ray. Furthermore, the binding entropy also correlates with the alkyl substituents having a more favorable desolvation entropy.

Temperature dependent studies ( $\Delta C_p$ ) on **3.42**, a representative alkyl ligand, suggests that more apolar surface area was buried upon complexation of **3.42** to the Grb2-SH2 domain than our benchmark ligand **3.04**. This suggested that the favorable entropy observed upon binding arose from entropy driven hydrophobic effects.

Several studies have surfaced recently disputing the claim that addition of apolar surface area leads to affinity enhancements in protein ligand interactions that are *driven by enthalpy not entropy*.<sup>6,9</sup> Thus, coming into this study we asked the question: How will adding apolar surface area to ligands affect the binding thermodynamics of protein-ligand complexation? Our findings presented herein do confirm that addition of ligand apolar surface area can be accompanied by a favorable change in *enthalpy*, but we also offer a revision: Although we have confirmed that the addition of apolar surface area to known Grb2 binders led to enthalpically driven binding increases, we have shown that this effect may dependent upon the surface topology lipophilic group. Hence, the more rigid, aromatic substituents **3.08**, **3.13** and **3.39** that bound to Grb2-SH2 domain consistently bound with more favorable binding *enthalpies* relative to both the alkyl ligands (**3.42–3.44**) and a common benchmark (**3.04**). Presumably this is due to more favorable vdW contacts that the aromatic ligands make in the binding pocket, although we were only able to obtain limited structural data to support this hypothesis. Therefore, our study does suggests that the increase of lipophilicity of

a ligand can increase the ligand potency through favorable binding enthalpies, but this process can also be driven by entropy depending on the ligand surface topology. This indicates that there still may be unknown components to increasing the hydrophobicity of protein-binding ligands that are worthy of further investigation

## Chapter 4 Topliss Approach to the Binding Affinity Enhancement of MUP-I Binders.

### 4.1 Introduction

One of the underlying challenges in the discovery of new medicinal drugs is the design and synthesis of small organic molecules that bind tightly to biological receptors. Despite many recent advances in the area of *in silico* drug design,<sup>clxxii,clxxiii,clxxiv,clxxv,clxxvi</sup> computer scoring functions are still not capable of reliably calculating binding affinity from molecular structure. As a result, several strategies have been developed to optimize binding affinity of ligands. One of the earlier methods that aims to streamline the traditional optimization of drug leads are the operational schemes of the Topliss decision tree, whereby a small set of initial analogs are synthesized, and their measured binding affinities then guide the synthesis of an optimal compound.<sup>clxxvii,clxxviii</sup> These choices in ligand structure are dictated by considerations of hydrophobicity, steric effects, and electrostatic factors of differing substituents.

Although the Topliss operational schemes were originally published in 1972, they are frequently employed in contemporary drug design and discovery.<sup>clxxvii,clxxviii</sup> As of the submission of this dissertation, at least 40 unique drug discovery projects in both the private and public sector have employed these schemes within the past five years.<sup>clxxix</sup> However, there has yet to be a single study that examined the detailed thermodynamic effects associated with applying the

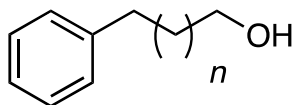
Topliss tree. This is very surprising considering the fact that the Topliss schemes are intended first and foremost to identify useful *trends* for the synthesis of potent drug leads.<sup>clxxx</sup> Thus, we hoped that in analyzing the thermodynamic trends associated with the Topliss operational schemes, we could discover ways in which they could be used more efficiently.

## 4.2 Experimental Design

We identified a set of phenyl alkanols ranging from 3-phenylpropanol to 6-phenylhexanol that bound to MUP-I with moderate to very good binding affinities (Table 4.1).<sup>clxxxi</sup> It was observed that the ligand affinity increased steadily upon the addition of methylene groups from  $n = 1$  until  $n = 3$ , at which point the affinity reached its zenith. For the six carbon analog ( $n = 4$ ), binding affinity decreased slightly. This was attributed to an unfavorable syn-pentane interaction that the ligand adopts in the binding pocket that is visible in the crystal structure 6-phenylhexanol **4.00** complexed with MUP-I (figure 4.1).



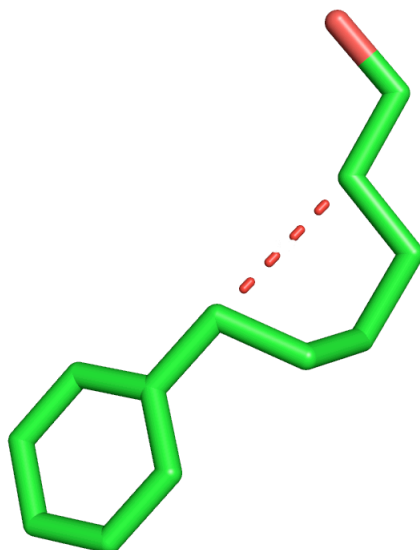
**Table 4.2.1** Binding of phenyl alkanols with MUP-I at 298 K in PBS.<sup>a</sup>



<i>n</i>	$K_a$ ( $\times 10^5 \text{ M}^{-1}$ )	$\Delta G_{obs}^\circ$ (kcal•mol <sup>-1</sup> )	$\Delta H_{obs}^\circ$ (kcal•mol <sup>-1</sup> )	$-T\Delta S_{obs}^\circ$ (kcal•mol <sup>-1</sup> )
<b>1</b>	<b>1.1 ± 0.03</b>	<b>-6.9 ± 0.01</b>	<b>-12.2 ± 0.47</b>	<b>5.3 ± 0.1</b>
<b>2</b>	<b>3.0 ± 0.03</b>	<b>-7.4 ± 0.02</b>	<b>-13.9 ± 0.69</b>	<b>6.4 ± 0.1</b>
<b>3</b>	<b>11.1 ± 0.04</b>	<b>-8.3 ± 0.07</b>	<b>-16.5 ± 0.71</b>	<b>8.2 ± 0.3</b>
<b>4</b>	<b>10.0 ± 0.05</b>	<b>-8.2 ± 0.2</b>	<b>-13.3 ± 0.60</b>	<b>5.1 ± 0.6</b>

<sup>a</sup>ITC experiments were conducted at 25 °C in PBS buffer at pH 7.45 ± 0.05. Errors based on 5% error in ligand concentration plus standard deviations.

**Figure 4.2.1** Pose of 6-phenylhexanol (**4.00**) (Table 4.1,  $n = 4$ ) in the binding complex of MUP-I

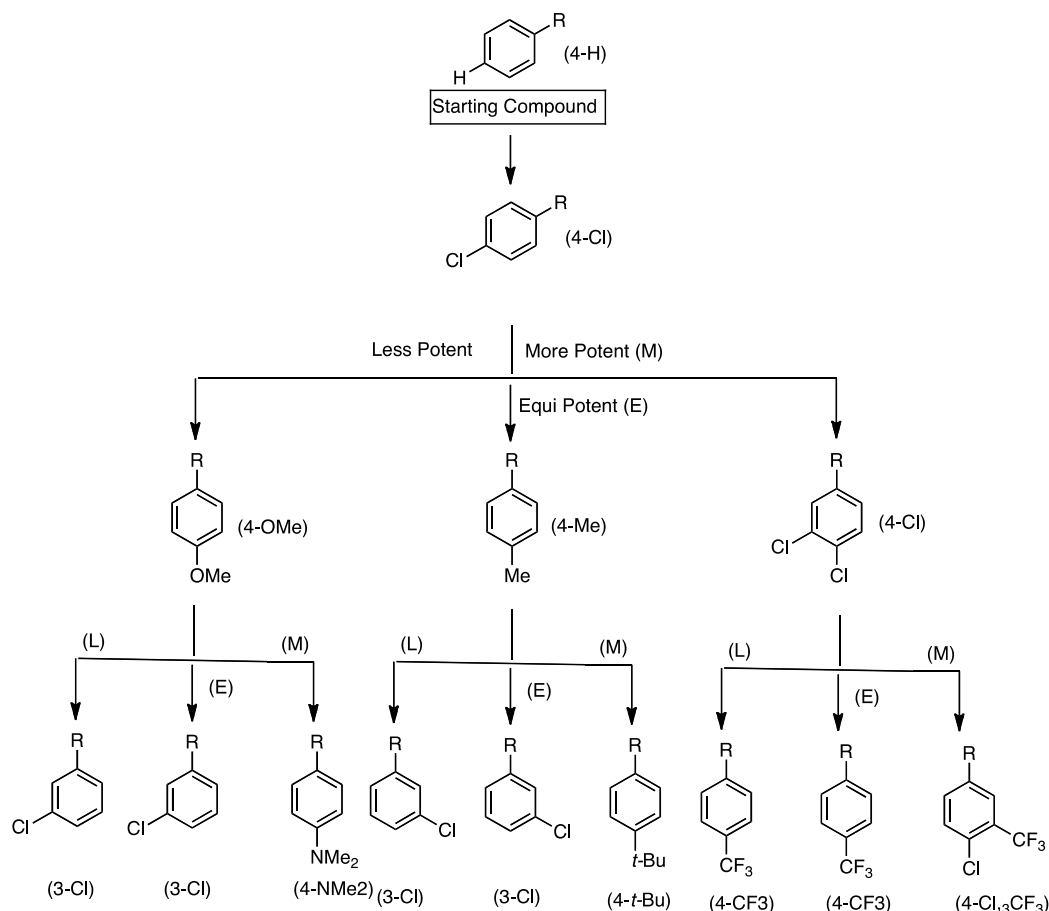


Several features of the phenyl alkanol that we had investigated in Chapter 2 immediately intrigued us. First, since these ligands contained an unsubstituted phenyl ring, we quickly envisioned the utility of this carbon platform skeleton as a model system for studies utilizing the Topliss decision tree. Second, because we had previously discovered that the binding affinity of these ligands was effected by the addition/subtraction of methylene groups, we were also hopeful that we would be able to obtain differences in the ligand binding affinity that were not severely compromised from enthalpy-entropy compensation.<sup>clxxxii</sup> Lastly, we felt that MUP-I would serve as an appropriate model system for our Topliss studies because its binding associations with small molecules are postulated to be driven by nondirectional van der Waals (vdW) contacts,<sup>clxxxiii</sup> and as such there are no

obvious polar networks that we would need to maintain. Thus, we initiated our studies with phenyl butanol since it showed the median bioactivity of the four ligands shown in Table 4.2.1. Using phenyl butanol as a starting point seemed likely it to maximize our chances of observing affinity enhancements, and it would also provide us the opportunity to apply the schemes to ligands that were nearly an order of magnitude more or less potent.

Shown in *Figure 4.2.2* is the classic Topliss decision tree that was reported in the original publication.<sup>11</sup> The strategy requires one to first synthesize a 4-Cl derivative of an initial compound and then test its bioactivity. The chart then guides one to the next compound based on whether or not the analog binds with more, equal or less potency than the previous iteration. This process is repeated until the potency of the initial hit increases.

**Figure 4.2.2** Topliss decision tree from an unsubstituted phenyl ring of a known drug lead

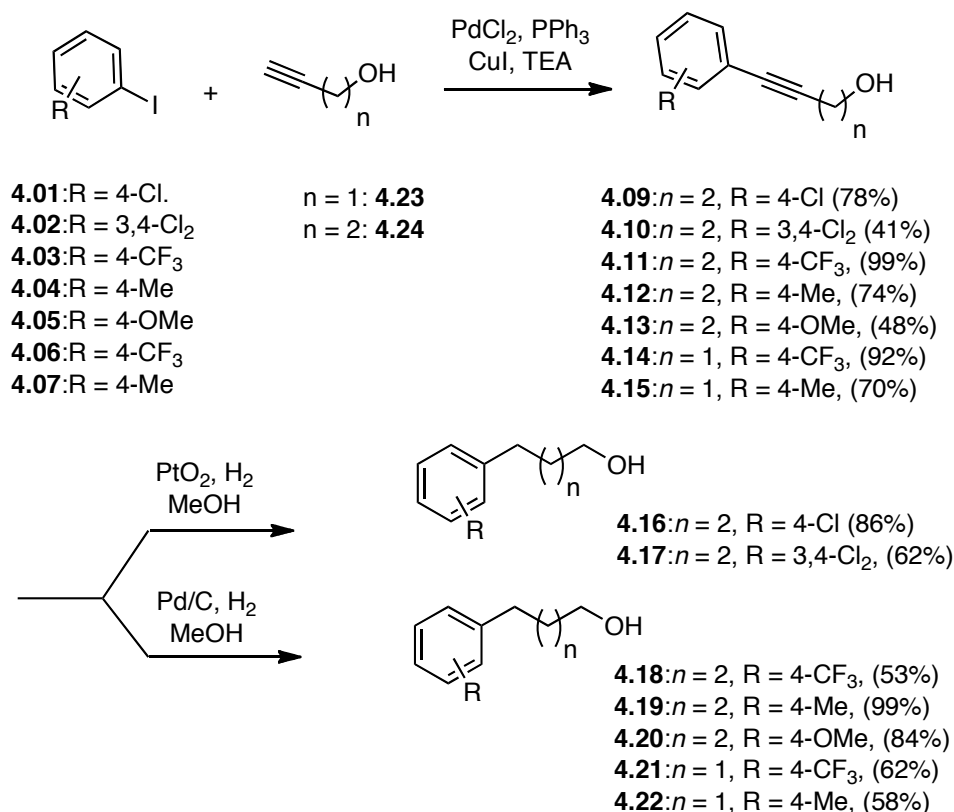


### 4.3 Ligand Synthesis

All of the alcohols that were targeted in this study were readily synthesized *via* a Sonogashira coupling from the related aryl iodides. Accordingly, the reaction of aryl halides **4.01-4.07** with alkyne **4.08** in the presence of PdCl<sub>2</sub> and CuI provided the requisite alkynes in 42 – 91% yields. To obtain the saturated alcohols **4.18-4.22**, the alkyne was reduced by palladium-catalyzed hydrogenation. Because these conditions led to dehalogenation of ligands **4.16** and **4.17**, the ligands were prepared by reduction using PtO<sub>2</sub>, which is known to be less likely to hydrogenolyze C-X bonds (scheme 4.3.1). For the discussions of

the binding and thermodynamic properties of **4.16-4.22**, we developed a code system based on their substitution pattern such that **4.18** is written as **C4pCF<sub>3</sub>**, such methylene carbons are coded as “C” followed by the number of methylene groups in the ligand, followed by the phenyl substitution pattern.

**scheme 4.3.1** Ligand synthesis of desired Topliss derived alkanols.

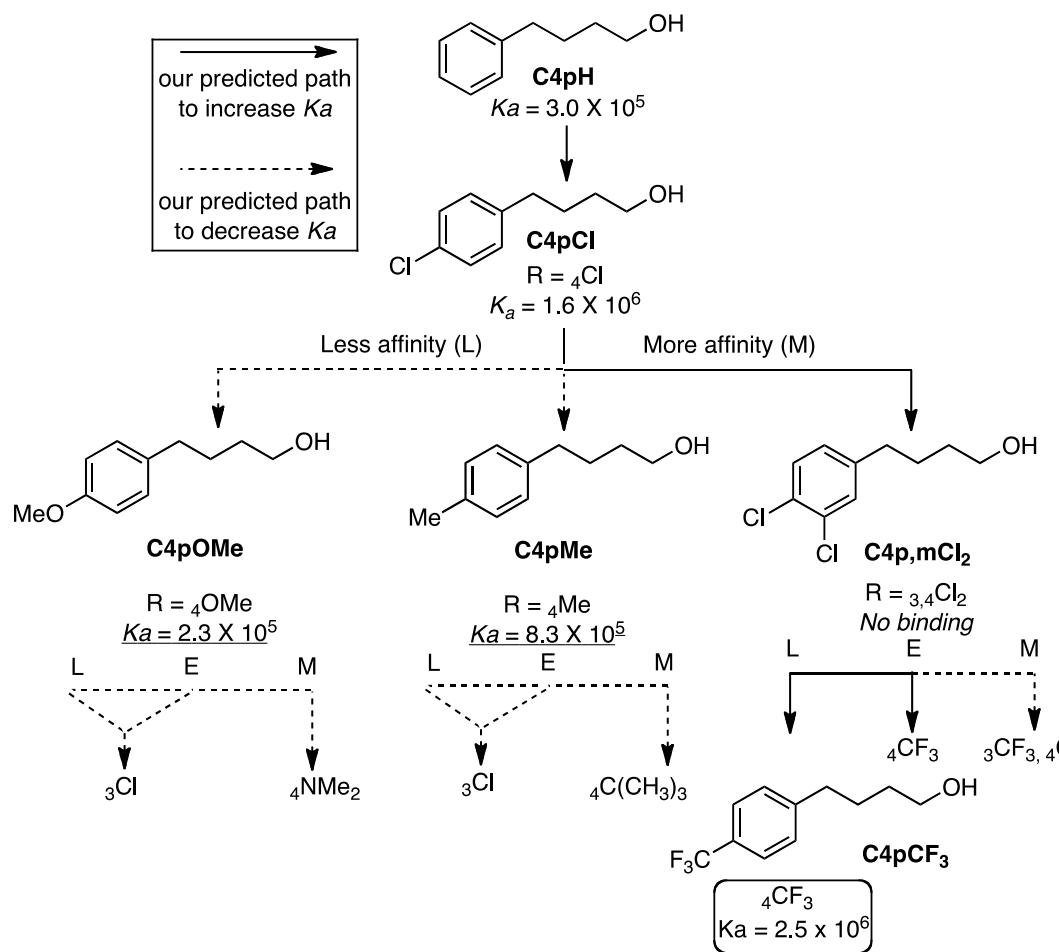


We determined the binding energetics of these protein-ligand interactions using ITC. To our delight, the affinity of the initial analog **C4pH** to MUP-I was expediently optimized following strict adherence to the Topliss operational schemes (Figure 4.3.1). In the manner that the charts are meant to reveal the nature of substituents that favor and disfavor binding, we quickly learned what

changes in the substituents led to both increased and decreased ligand potency. For example, the notable increase in the binding affinity with the 4-Cl substitution that was not observed with the 4-OMe indicates that there may be a preference in this system for electron withdrawing over electron donating substituents. Furthermore, a slight increase in binding affinity was observed when the electron donating capabilities of the ligand were reduced as evident by comparing 4-OMe **C4pOMe** to the 4-Me analog **C4pMe**. This increase in potency may also be attributed to the removal of a polar atom from the ligand. The drop in binding affinity for the 3,4-Cl<sub>2</sub> **C4p,mCl<sub>2</sub>** was initially attributed to added steric encumbrance in the pocket, but no structural data could be obtained to support this hypothesis.

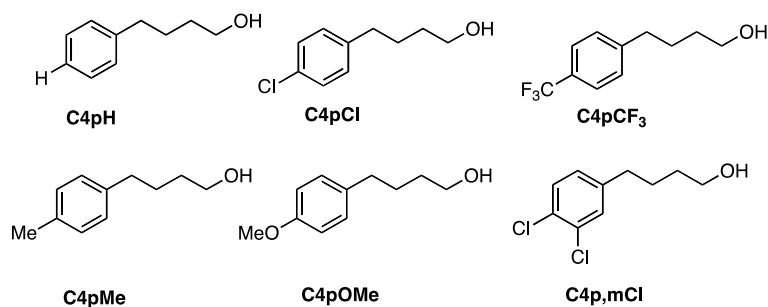
Finally, the schemes correctly led us to the 4-CF<sub>3</sub> substituted ligand **C4pCF<sub>3</sub>** the most potent analog. That both the 4-OMe and 4-Me derivatives did not provide significant affinity enhancements validates the utility of these schemes for this system because they were not predicted provide affinity boosts based off of our testing of the initial 4-Cl analog.

**Figure 4.3.1.** Topliss Analysis.



We next analyzed the thermodynamics of substituent effects that led to affinity enhancement (Table 4.3.1). In general, we found that all the substituents that led to an increase in binding affinity had a *more* favorable entropy of binding than the starting unsubstituted ligand (shown in Graph 4.1). Therefore, the general trend in binding entropy that we observe between the Topliss derivatives and the parent analog also correlates well with the increase in the Hansch- $\pi$  values<sup>clxxxiv</sup> becoming more positive ( $R^2 = 0.96$ , Graph 4.3.2).

**Table 4.3.1** ITC analysis of Topliss-derived analogs for MUP-I at pH 7.4 in PBS.<sup>a</sup>



Ligand	$K_a$ ( $\times 10^5 \text{ M}^{-1}$ )	$\Delta G_{obs}^\circ$ ( $\text{kcal}\cdot\text{mol}^{-1}$ )	$\Delta H_{obs}^\circ$ ( $\text{kcal}\cdot\text{mol}^{-1}$ )	$-T\Delta S_{obs}^\circ$ ( $\text{kcal}\cdot\text{mol}^{-1}$ )	$\pi$
<b>C4pH</b>	<b><math>3.0 \pm 0.03</math></b>	<b><math>-7.4 \pm 0.02</math></b>	<b><math>-13.9 \pm 0.7</math></b>	<b><math>6.4 \pm 0.1</math></b>	0
<b>C4pOMe</b>	<b><math>4.3 \pm 0.01</math></b>	<b><math>-7.7 \pm 0.01</math></b>	<b><math>-14.4 \pm 0.6</math></b>	<b><math>6.7 \pm 0.3</math></b>	-0.02
<b>C4pCl</b>	<b><math>16.0 \pm 0.1</math></b>	<b><math>-8.4 \pm 0.01</math></b>	<b><math>-12.2 \pm 0.5</math></b>	<b><math>3.8 \pm 0.2</math></b>	0.71
<b>C4pMe</b>	<b><math>8.4 \pm 0.02</math></b>	<b><math>-8.2 \pm 0.01</math></b>	<b><math>-11.4 \pm 0.5</math></b>	<b><math>3.2 \pm 0.3</math></b>	0.56
<b>C4pCF<sub>3</sub></b>	<b><math>28.0 \pm 0.2</math></b>	<b><math>-8.8 \pm 0.05</math></b>	<b><math>-14.0 \pm 0.4</math></b>	<b><math>5.2 \pm 0.2</math></b>	0.88
<b>C4p,mCl<sub>2</sub></b>	<b>Did not Bind (<math>K_a &lt; 1 \times 10^4</math>)</b>				1.25

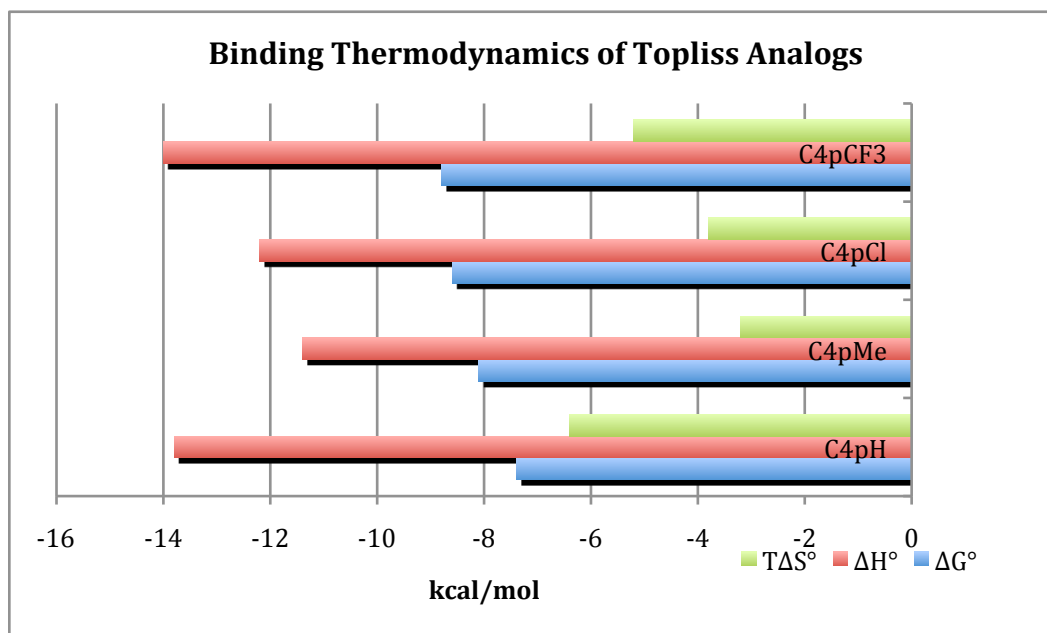
<sup>a</sup>ITC experiments were conducted at 25 °C in PBS buffer at pH  $7.45 \pm 0.05$ . Errors based on 5% error in ligand concentration plus standard deviations. Ligands are referred to by their phenyl substituent and alkyl linker length, such that 4-(4-(trifluoromethyl)phenyl)butanol (**4.18**) is listed as **C4pCF<sub>3</sub>**.

Comparing the binding energetics for the pair **C4pH** and **C4pCl** was reveals that **C4pCl** binds with a more favorable affinity because of an increase in binding entropy of  $2.6 \text{ kcal}\cdot\text{mol}^{-1}$ . There is a small increase in affinity from **C4pCl** to **C4pCF<sub>3</sub>** which is due to a favorable  $\Delta\Delta H^\circ$  of  $1.8 \text{ kcal}\cdot\text{mol}^{-1}$  that is compensated by an unfavorable  $T\Delta\Delta S^\circ$  of  $1.4 \text{ kcal}\cdot\text{mol}^{-1}$ . Conversely, for the ligand pair **C4pCl** and **C4pOMe**, the binding energetics reveal **C4pOMe** binds

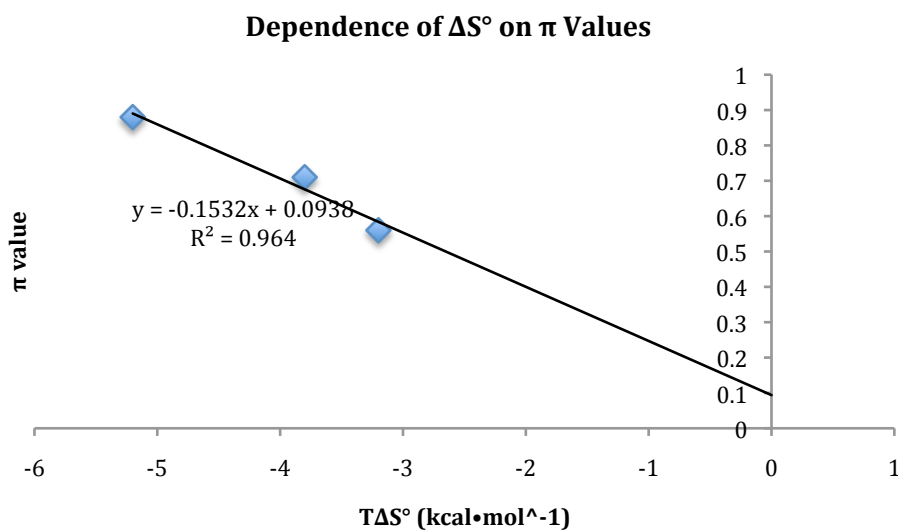


with a less favorable affinity because of a less favorable  $\Delta\Delta S^\circ$  of 2.9 kcal•mol<sup>-1</sup> that dominates a favorable  $\Delta\Delta H^\circ$  of 2.2 kcal•mol<sup>-1</sup>. Lastly, for the ligand pair **C4pCl** to **C4pOMe**, we observe that C4pOMe binds with a lower affinity because an unfavorable  $\Delta\Delta H^\circ$  of 2.9 kcal•mol<sup>-1</sup> dominates a favorable  $\Delta\Delta H^\circ$  2.2 kcal•mol<sup>-1</sup>. The substitution to the 3,4-Cl<sub>2</sub> analog led to a large decrease in binding enthalpy such that it could not be detected by our instrumentation. Typically, ligands that bind to a protein with an enthalpy so unfavorable that accurate thermographs cannot be obtained are thought to have a  $K_a < 1 \times 10^3 \text{ M}^{-1}$  as indicated in **Table 4.3.1**.

**Graph 4.3.1** Bar graph showing thermodynamic parameters of binding of **C4pH**, **C4pMe**, and **C4pCl** and **C4pCF3** to MUP-I.



**Graph 4.3.2** Graph of  $\pi$  value correlation to  $T\Delta S^\circ$



It was also notable that the effect of replacing a methyl group with a trifluoro group was accompanied by a favorable change in relative binding that was partially offset by an unfavorable change in relative binding entropy of  $2.0 \text{ kcal}\cdot\text{mol}^{-1}$  that led to an increase in

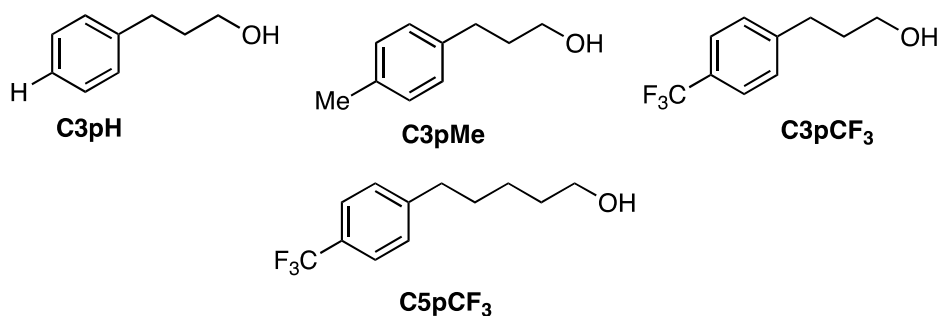
affinity of 0.7 kcal•mol<sup>-1</sup>. It is worth noting that this observation is not consistent with the conventional view that C-H atom pair allows for better dispersive interactions in the binding pocket than the respective C-X atom pair. However, in a recent report by Whitesides,<sup>cxxxvii</sup> it was shown that ligands having C-F bonds consistently bound with higher enthalpy and affinity than in the respective C-H analogs. It is also worth noting that he also observed a favorable change in entropy with this ligand modification, which is not observed in our system.

#### 4.4 Effect of Fluorination of Alkyl Bonds.

It has been previously shown that adding methylene groups to a ligand is a viable strategy to enhance binding affinity in select MUP-I binders.<sup>clxxxv,clxxxvi,clxxxvii</sup> During the course of our own studies with the phenyl alkanols, we observed a very similar phenomenon whereupon added apolar surface area in the form of a methylene group led to *enthalpy* driven affinity enhancements (Chapter 2). Since we had already observed pronounced affinity boosts upon the addition of apolar surface area to the phenyl alkanols, we decided to probe the effects that we observed upon the introduction of C-F bonds on different derivatives of these binders (Table 4.2.1) wherein the length of the methylene linker was varied accordingly. This strategy allows us to investigate two significant scientific questions regarding protein-ligand interactions: (1) is there any dependence on the initial affinity of the starting analog that effects the Topliss-derived affinity boosts? And (2), since both methylene addition *and* Topliss substituent effects are shown to grant affinity increases in this system, can we combine these two modifications to obtain synergistic, additive affinity gains?<sup>clxxxviii,clxxxix,cxc</sup>

Inspired by the high affinity obtained by 4-CF<sub>3</sub> substitution, we first synthesized and tested the **C5pCF<sub>3</sub>** analog of our initial study, only to find that the affinity of this ligand for MUP-I was below the detection limit of  $1 \times 10^3 \text{ M}^{-1}$  of the instrument (Table 4.4.1). The decline in binding affinity was not surprising because extending the methylene linker of our phenyl alcohols eventually led to deleterious effects in binding affinity (Table 4.2.1). Following this result, we next synthesized derivatives of our phenyl alkanols wherein  $n = 1$  (Table 4.4.1). As in the case where  $n = 2$ , The increase in affinity of **C3pMe** to **C3pCF<sub>3</sub>** was again accompanied by a very favorable enthalpy of  $1.6 \text{ kcal}\cdot\text{mol}^{-1}$  that compensated an unfavorable change entropy of  $0.8 \text{ kcal}\cdot\text{mol}^{-1}$  such that binding affinity was more favorable by  $0.7 \text{ kcal}\cdot\text{mol}^{-1}$ . in which correlates very well to the trend that we observed for the phenylbutanol system.

**Table 4.4.1.** ITC of expanded series in PBS at pH 7.4 at 298 K.<sup>a</sup>



	$K_a$	$\Delta G_{obs}^\circ$	$\Delta H_{obs}^\circ$	$-T\Delta S_{obs}^\circ$
Ligand	( $\times 10^5 \text{ M}^{-1}$ )	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )
<b>C3pH</b>	<b>1.1 ± 0.03</b>	<b>-6.9 ± 0.01</b>	<b>-12.2 ± 0.5</b>	<b>5.3 ± 0.1</b>
<b>C3pMe</b>	<b>8.8 ± 0.03</b>	<b>-8.2 ± 0.01</b>	<b>-13.7 ± 0.5</b>	<b>5.6 ± 0.1</b>
<b>C3pCF<sub>3</sub></b>	<b>34.1 ± 0.3</b>	<b>-8.9 ± 0.04</b>	<b>-15.3 ± 0.4</b>	<b>6.4 ± 0.2</b>
<b>C5pCF<sub>3</sub></b>	<b>Did not Bind (<math>K_a &lt; 1 \times 10^4</math>)</b>			

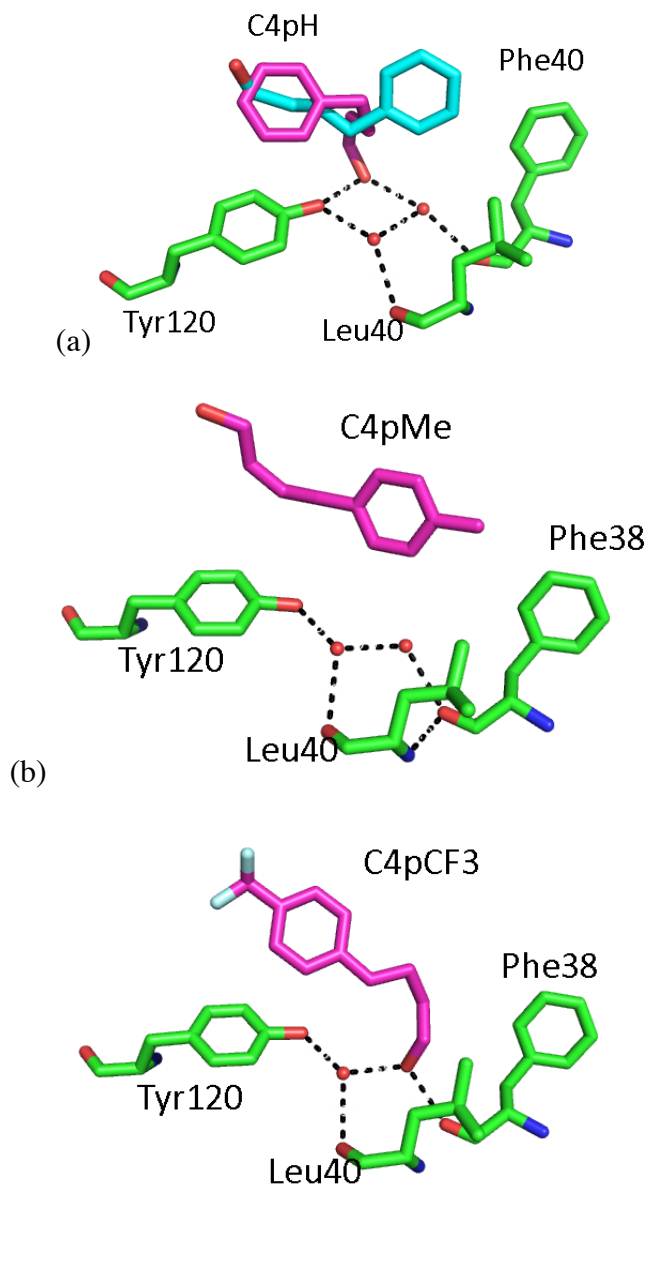
<sup>a</sup>ITC experiments were conducted at 25 °C in PBS buffer at pH 7.45 ± 0.05. Errors based on 5% error in ligand concentration plus standard deviations. Ligands are referred to by their phenyl substituent and alkyl linker length, such that 3-(4-(trifluoromethyl)phenyl)propan-1-ol (**4.21**) is listed as **C3pCF<sub>3</sub>**.

## 4.5 Structural Characterization

X-ray crystal structures of MUP-I complexed with ligands **C4pH**, **C4pMe**, and **C4pCF<sub>3</sub>** were obtained by Dr. John Clements to a resolution of <2.0 Å (Figure 4.5.1). The number of crystallographic water molecules and total number of vdW contacts in the binding pocket of the complexes were tallied. Backbone atoms belonging to the protein in the complexes align closely (RMSD < 0.2 Å). However, the position and conformation of the bound ligands in the pocket vary considerably from one complex to another. For example, **C4pH** binds to MUP-I in two different poses, such that in one pose the

hydroxyl group on the ligand makes seven direct polar contacts with the phenolic oxygen atom of Tyr120 and the other it makes two. Polar protein-ligand networks involving direct and water-mediated contacts to protein residues Phe38, Leu40, and Tyr120 were also observed for the analogs bound to MUP-I (**Figure 4.5.1**). For ligand complexes with **C4pH** and **C4pMe**, there are two bound water molecules in the complex. However, for the ligand complex of **C4pCF3**, there is only one bound molecule in the water complex.

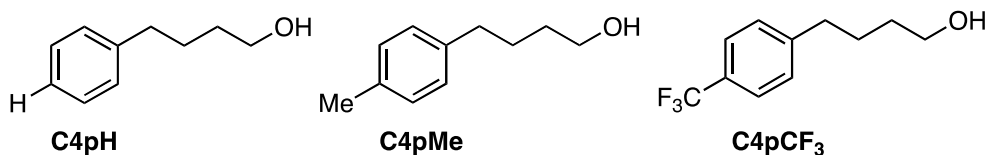
**Figure 4.5.1** Structural complexes of MUP-I with the Topliss derived analogs **C4pH**, **C4pMe**, and **C4pCF<sub>3</sub>** were obtained to a resolution of <2.0 Å research by researcher Dr. John Clements.<sup>a</sup>



<sup>a</sup> Oxygen and nitrogen atoms are colored red and blue, respectively. Carbon atoms belonging to the protein (line) are colored green while those belonging to the ligands (sticks) are colored magenta or cyan, the latter shown for complexes in which the bound ligands occupy more than one pose. Protein-ligand direct contacts and those mediated by water molecules (red spheres) are indicated with dashed, black lines. (a) Complex with **C4pH**. (b) Complex with **C4pMe**. (c) Complex with **C4pCF<sub>3</sub>**

The vdW contacts made in the binding complex were tallied and are shown in Table 4.5.1 As previously noted, the **C4pH** analog binds in two different poses in the binding pocket, with one making 49 vdW contacts in one pose and 33 in the other. This observation highlights one of the difficulties with drawing conclusions from counting vdW contacts, because two isoenergetic protein-ligand complexes show some variation. However, the number of vdW contacts observed for **C4pMe** and **C4pCF<sub>3</sub>** show less variation: **C4pMe** makes 42 contacts and **C4pCF<sub>3</sub>** makes 50 contacts. Since this falls within the range that we observed for **C4pH**, it is difficult to assign the significance of any differences in the number of contacts to  $\Delta\Delta G^\circ$ ,  $\Delta\Delta H^\circ$ , or  $T\Delta\Delta S^\circ$ . Thus, there are no correlations of thermodynamics parameters with the number of vdW contacts.

**Table 4.5.1** Structural characteristic that were observed from the protein bound complexes of **C4pH**, **C4pMe** and **C4pCF<sub>3</sub>**.



	$\Delta G_{obs}^\circ$	$\Delta H_{obs}^\circ$	$-T\Delta S_{obs}^\circ$	vdW	Buried $\Delta$ CSA
Ligand	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )	(kcal•mol <sup>-1</sup> )	contacts	(Å <sup>2</sup> )
<b>C4pH</b>	<b>-7.4 ± 0.02</b>	<b>-13.9 ± 0.7</b>	<b>6.4 ± 0.1</b>	<b>49,33</b>	<b>153.6, 154.7</b>
<b>C4pMe</b>	<b>-8.2 ± 0.01</b>	<b>-11.4 ± 0.5</b>	<b>3.2 ± 0.3</b>	<b>42</b>	<b>165.3</b>
<b>C4pCF<sub>3</sub></b>	<b>-8.8 ± 0.05</b>	<b>-14.0 ± 0.4</b>	<b>15.2 ± 0.2</b>	<b>50</b>	<b>174.6</b>

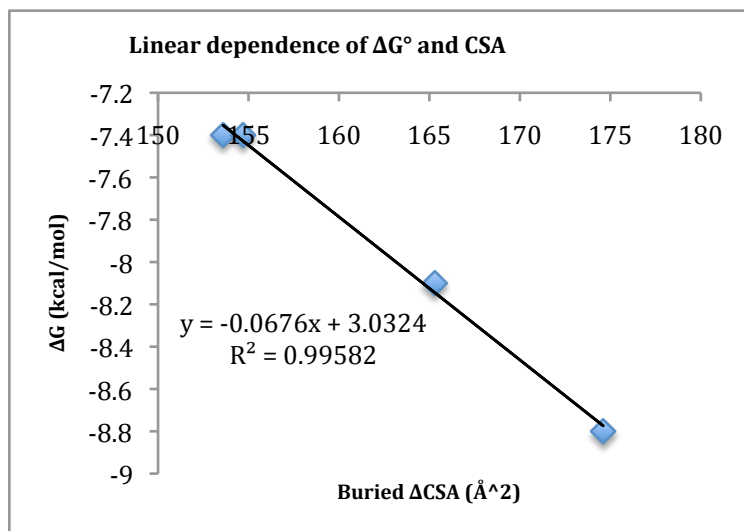
We then compared the differences in  $\Delta G^\circ$  with ligand Connolly surface area<sup>exci</sup> ( $\Delta$ CSA) that occurred upon binding for the three ligands for which we had obtained



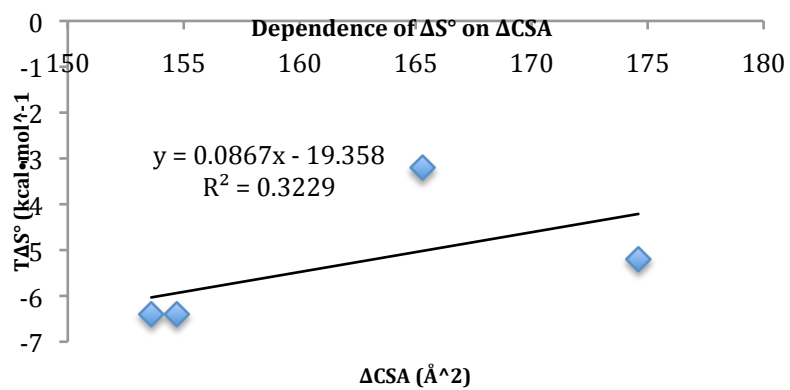
cyrstallographic data. The  $\Delta\text{CSA}$  correlates strongly with  $\Delta G^\circ_{\text{obs}}$  for these ligands ( $R^2 = 0.99$ ) with a slope of  $-68 \text{ cal/mol}\text{\AA}^{-2}$  (Graph 4.5.1). That the burial of nonpolar surface area is consistent with *more favorable* entropies of binding relative to **C4pH** is suggestive of an entropically driven hydrophobic effect; however, the correlations to  $\Delta H^\circ$  and  $T\Delta S^\circ$  are not very strong.

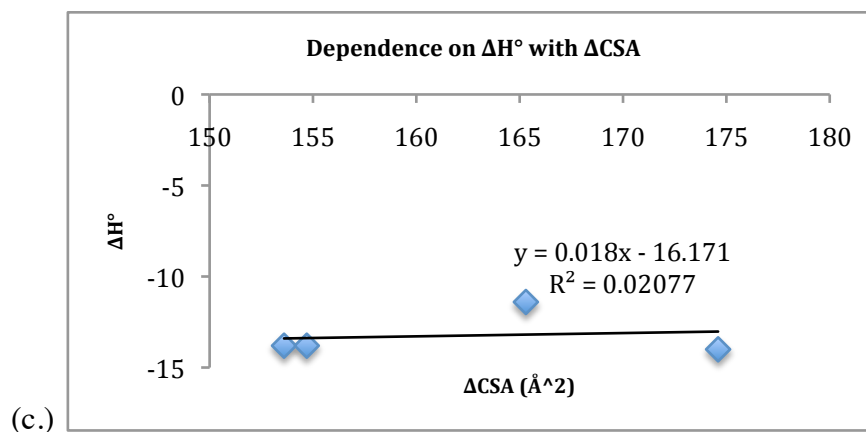
**Graph 4.5.1** Dependence of  $\Delta\text{CSA}$  on (a)  $\Delta G^\circ$ , (b)  $\Delta H^\circ$ , and (c)  $T\Delta S^\circ$ .

(a).



(b.)





That the optimization of phenylbutanol using the Topliss approach led to favorable increases in entropy was intriguing. Typically with MUP-I binders, affinity is increased through the addition of nonpolar surface area that imparts more favorable binding enthalpy changes.<sup>clxxxv,clxxxvi</sup> However, increases in entropy were observed as the dominant driving force in the Topliss studies shown in Figure 4.3.1. Furthermore, the crystal data indicated that the increases in binding affinity of our initial ligand correlated very well with  $\Delta\text{CSA}$ . Such entropically driven increases in binding affinity are considered to arise from hydrophobic effects.

Spaller and coworkers recently studied the effects of introducing halogen atoms on the aromatic an aromatic ring of compounds that bind to the PDZ domain.<sup>excii</sup> They also found that halogen substitution on phenyl rings increased binding affinity through an *entropically* favorable process. This change in entropy was attributed to the increase in the vdW radii of the halogens relative to an H atom at the *para* position of a phenyl ring (H: 1.20 Å, F: 1.47 Å, Cl: 1.77 Å).<sup>exciii</sup> This trend is a similar to what we observed upon the substitution of a chloro group on the phenyl ring of **C4pH**. Thus, it could be that the

favorable entropic gains arise from ligand desolvation, however, we observe a very poor correlation between  $\Delta\text{CSA}$  and  $\Delta S^\circ$ . It is likely that other components are at play.

A similar observation by the Whitesides group in a study on carbonic anhydrase II (HCA II) binder was postulated to arise from differences in dehydration of the binding pocket between the C-H and C-F groups, since the dispersive interactions in the complex would be weaker for C-F bonds than they are in C-H groups.<sup>Error! Bookmark not defined.</sup> It was speculated that the larger radii of the fluoroalkyl group may displace more unoptimally hydrated water molecules that would then be hydrated in bulk water. It is thought that one face of the active site in HCA II is a hydrophobic “wall” and therefore may contain unoptimally hydrated water molecules.<sup>cxciiv</sup> Upon obtaining crystal structures of ligands **C4pMe** and **C4pCF<sub>3</sub>**, we observed that there was one less water molecule in the ligand complex of **C4pCF<sub>3</sub>**. However, one must be careful when extrapolating meaningful correlations from crystallographic data, but it is at least noteworthy that the displacement of a water molecule does correlate with a favorable change in enthalpy that we observe upon modifying alkyl groups to fluoroalkyl groups, which arises from the water molecule making more *enthalpically* favorable hydrogen bond networks in bulk solvent.

#### 4.6 A possible dependence of the Topliss schemes on initial analog potency.

Based off of the results from expanding the effects of the fluorine groups to phenyl alkanols of different initial potency, we became interested in studying if there was an effect of the affinity gained by the Topliss-derived modifications and the starting potency of the initial lead compound. To analyze this concept further, we decided to rank our Topliss derived ligands such that their affinity boost is calculated relative to their initial potency. Therefore, if we are to assume for *this study* that the ligand panel that was

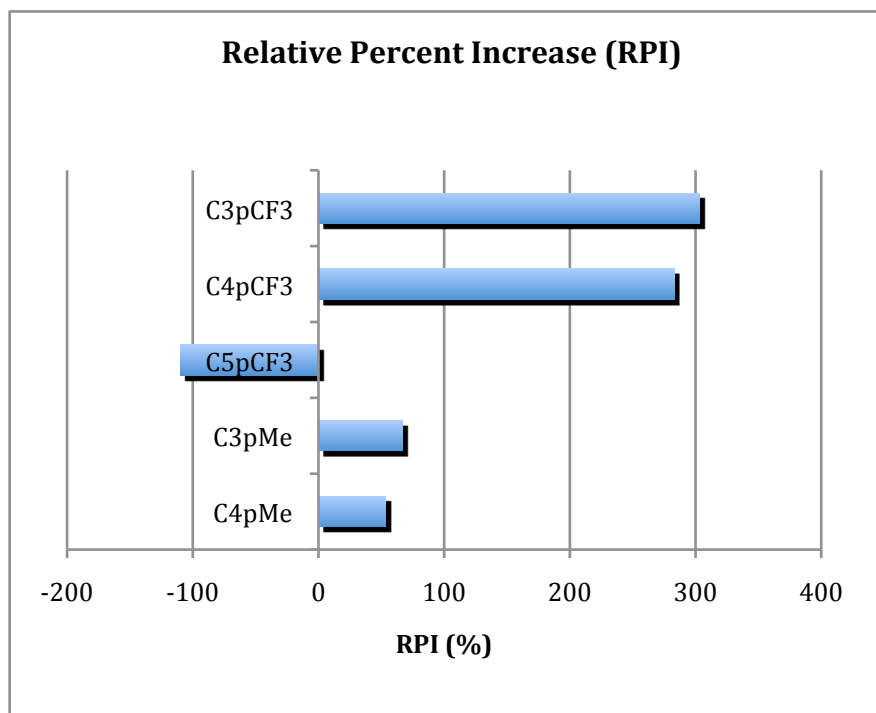
shown in Table 4.2.1 represents the spectrum of binding for MUP binders, then phenylpentanol would represent the theoretical *ceiling* for potency, and phenylpropanol would represent the optimization *floor*. For the sake of this model system shown in Figure 4.2.1, the difference in affinity of these two ligands can be referred to as the optimization index. Any optimization of a weak ligand could be observed relative to the optimization index to ascertain the magnitude of affinity enhancement that is being gained (or lost) by the chemical transformation on a starting ligand, this may be called the relative percent increase (RPI) of the newly tested analog (eq. 4.1). For example, the RPI of an analog (C) of starting ligand A, could be calculated as follows:

$$\text{RPI} = \frac{K_{a(C-A)}}{\text{Optimization Index}} \times 100\% \quad (\text{eq. 4.1})$$

When we look at the RPI of the **C4pMe** and **C4pCF3** analogs used in the Topliss decision tree (Figure 4.3.1), it is observed that there is a greater affinity boost in all cases where the structural modification was applied to the weakest of the phenyl alcohols in the original panel (Figure 4.6.1 Table 4.2.1 ). Therefore, although the Topliss tree did provide direction towards a higher affinity binder, which was one of our goals, we observed that it was more effective when applied to weaker ligands and that its efficacy tapered down as it was applied to more potent binders. This analysis suggests that the Topliss schemes may be less effective for advanced intermediates that bind with greater affinity. However, further ligands would be necessary to bulk up this claim. Therefore, future studies would be to synthesize **C3pCl**, **C3pOMe** and **C3pCl<sub>2</sub>** and test their binding affinities. Furthermore, one could also envision extending this study out even further to phenylehtnaol, which was shown in our labs to have a binding affinity that was slightly

less than instrumentation limit of  $K_a = 1 \times 10^3 \text{ M}^{-1}$ , which is to say that there was evidence that it bound to MUP-I, but unfortunately the  $\Delta H^\circ$  was too weak to obtain reproducible data.

**Figure 4.6.1** Increase in the binding affinity of each analog relative to the known optimal increase at the onset of study. Herein called the relative percent increase or RPI.



## Conclusion

We have described a case study where application of the Topliss operational schemes led to an expedient development of novel MUP-I binding analogs. We used ITC to measure the binding affinities for the Topliss-derived analogues and then followed the suggested path of the Topliss decision tree to obtain significant affinity enhancements of our lead compound. These analogs were then further decomposed by studying the thermodynamics and structural properties of binding. It was discovered that the Topliss-derived analogs in general bound with more favorable *entropy* of binding relative to the

“lead” **C4pH**. Although no consistent correlations were observed between  $\Delta G^\circ$  and  $T\Delta S^\circ$ , we did observe a rather strong correlation between the  $T\Delta\Delta S^\circ$  and the relative  $\pi$  values that correspond to the ligand’s phenyl substitution. Furthermore, we also found a very strong correlation with  $\Delta G^\circ$  and  $\Delta\text{CSA}$  for the ligands for which we obtained crystal structures.

Analysis of the Topliss-derived ligands also led us to discover and explain in thorough detail a trend regarding the binding differential between fluoroalkyl and alkyl substituents. It was observed in two different cases that analogs with C-F bonds bound with more favorable changes in binding enthalpies. This is a significant discovery considering that many medicinally relevant molecules are fluorinated. It has been proposed by Whitesides that the driving force of this event is the displacement of unoptimally hydrated water molecules in the binding pocket due to the larger radii of the fluorine atom. Indeed, it was observed in our studies that the structural data that the ligand-bound complex of the fluorinated analog contained one less water molecule in the binding pocket than our “lead” **C4pH**. Although this is an intriguing result, we have herein observed two cases where fluoroalkyl ligands show more pronounced changes in binding enthalpy relative to their corresponding alkyl variant. Upon attempting to study this effect on **C5pH**, the ligand binding event was no longer detectable by our instrumentation, which based off of our previous studies with the phenylalkanols may be due to added steric encumbrance in the binding pocket. Unfortunately, no structural data was obtained to confirm this hypothesis. Thus, our analysis suggests that further studies on a modified ligand model systems could be designed to further probe the binding thermodynamics of fluoroalkyl chains with MUP-I.

## Chapter 5 Experimental

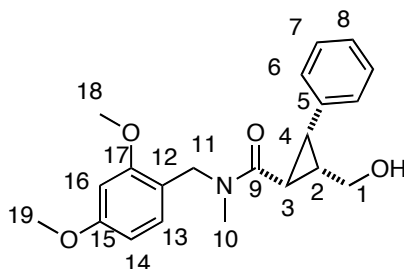
### 5.1 General

Tetrahydrofuran and diethyl ether were dried by filtration through two columns of activated, neutral alumina according to the procedure described by Grubbs.<sup>cxcv</sup> Methanol (MeOH), acetonitrile (MeCN), and dimethylformamide (DMF) were dried by filtration through two columns of activated molecular sieves, and toluene was dried by filtration through one column of activated, neutral alumina followed by one column of Q5 reactant. These solvents were determined to have less than 50 ppm H<sub>2</sub>O by Karl Fischer coulometric moisture analysis. Benzene, methylene chloride (CH<sub>2</sub>Cl<sub>2</sub>), diisopropylamine (*i*-Pr<sub>2</sub>NH), triethylamine (Et<sub>3</sub>N), diisopropylethylamine (*i*-Pr<sub>2</sub>Net), and pyridine were distilled from calcium hydride immediately prior to use. All reagents were reagent grade and used without purification unless otherwise noted, and air or moisture sensitive reagents were weighed in a glove box. All reactions involving air or moisture sensitive reagents or intermediates were performed under an inert atmosphere of nitrogen or argon in glassware that was flame or oven dried. Solutions were degassed using three freeze-pump-thaw cycles under vacuum. Reaction temperatures refer to the temperature of the cooling/heating bath. Volatile solvents were removed under reduced pressure using a Büchi rotary evaporator at 25–30 °C (bath temperature). Thin layer chromatography was run on pre-coated plates of silica gel with a 0.25 mm thickness containing 60F-254 indicator (EMD Millipore). Chromatography was performed using forced flow (flash chromatography) and the indicated solvent system on 230-400 mesh silica gel (Silicycle flash F60) according to the method of Still,<sup>cxcvi</sup> unless otherwise noted.



Infrared (IR) spectra were obtained either neat on sodium chloride or as solutions in the solvent indicated and reported as wavenumbers ( $\text{cm}^{-1}$ ). Proton nuclear magnetic resonance ( $^1\text{H}$  NMR) and carbon nuclear magnetic resonance ( $^{13}\text{C}$  NMR) spectra were obtained at the indicated field as solutions in  $\text{CDCl}_3$  unless otherwise indicated. Chemical shifts are referenced to the deuterated solvent (*e.g.*, for  $\text{CDCl}_3$ ,  $\delta = 7.26$  ppm and 77.0 ppm for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively) and are reported in parts per million (ppm,  $\delta$ ) relative to tetramethylsilane (TMS,  $\delta = 0.00$  ppm). Coupling constants ( $J$ ) are reported in Hz and the splitting abbreviations used are: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; comp, overlapping multiplets of magnetically nonequivalent protons; br, broad; app, apparent.

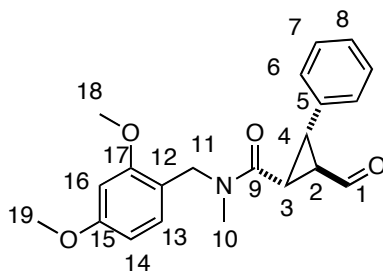
## 5.2 Procedures



**[1-R-(1 $\alpha$ -2 $\alpha$ -3 $\alpha$ )]-2-(4-Phenyl)-3-hydroxymethylcyclopropanecarboxylic acid-(2,4-dimethoxy-benzyl)-methylamide (2.19).** A freshly prepared solution of *tert*-BuMgCl in ether (2.0 M, 2.95 ml, 5.90 mmol) was added dropwise over 30 min to a solution of **2.16** (171 mg, 0.982 mmol) and **2.17** (485 mg, 2.95 mmol) in THF (2.1 mL) at  $-20$  °C. After stirring the mixture for 1 h at  $-20$  °C, the cooling bath was removed, and a saturated aqueous solution of  $\text{NH}_4\text{Cl}$  (5 mL) was added slowly with stirring. The mixture was extracted with  $\text{Et}_2\text{O}$  (3 x 10 mL), dried ( $\text{Na}_2\text{SO}_4$ ), filtered and concentrated under reduced pressure. The crude product was purified by silica gel chromatography eluting

with EtOAc:Hex (2:1) to yield 273 mg (78%) of the title product as a colorless oil.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ) (rotamers 1:1)  $\delta$  7.22 (d,  $J = 8.5$  Hz, 0.5 H), 7.13 (t,  $J = 11.3$  Hz, 1 H), 7.03 (br s, 0.5 H), 6.94 (d,  $J = 7.9$  Hz, 0.25 H), 6.47-6.45 (comp, 2 H), 5.12 (d,  $J = 8.2$  Hz, 0.5 H), 4.61 (comp, 1 H), 4.11 (d,  $J = 7.18$  Hz, 1 H), 3.83-3.78 (comp, 7 H), 3.67-3.63 (comp, 1 H), 3.32 (br s, 1 H), 3.06 (s, 1.5 H), 2.96 (s, 1.5 H), 2.63 (app t,  $J = 9.6$  Hz, 1 H), 2.27 (app t,  $J = 9.2$  Hz, 0.6 H), 1.91 (m, H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  171.4, 170.9, 160.9, 160.6, 158.8, 158.5, 136.3, 136.2, 131.6, 126.7, 126.5, 117.5, 116.4, 104.5, 104.3, 98.8, 98.4, 60.5, 59.3, 55.6, 55.5, 55.4, 49.3, 45.3, 35.8, 33.5, 31.1, 27.4, 27.1, 25.4, 25.3, 25.1, 24.1, 21.2; IR (DCM) 3400, 1613, 1456, 1415  $\text{cm}^{-1}$  mass spectrum (CI)  $m/z$  356.1861 [ $\text{C}_{21}\text{H}_{25}\text{NO}_4$  (M+1) requires 356.1860].

**NMR Assignments:** Rotamers (1:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.22 (d,  $J = 8.5$  Hz, 0.5 H, C13-H), 7.13 (t,  $J = 11.3$  Hz, 1 H, C8-H), 7.03 (br s, 0.5H, C6-H), 6.94 (d,  $J = 7.9$  Hz, 0.25 H, C7-H), 6.46-6.45 (comp, 2H, C15-H & C14-H), 5.1 (d,  $J = 8.2$  Hz, 0.47 H, C11-H), 4.61 (comp, 1.0 H, C11-H), 3.83-3.78 (comp, 7 H, C18-H & C17-H & C1-H), 3.32 (br s, 1 H, C1-OH), 3.06 (s, 1.5H, C10-H), 2.63 (app t,  $J = 9.6$  Hz, 1 H, C3-H), 2.27 (app t,  $J = 9.23$  Hz, 0.6 H, C4-H ), 1.91 (m, 1 H, C2-H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  171.4 (C9), 170.9 (C9), 160.9 (C15 or C19), 160.6 (C15 or C19), 158.8 (C15 or C19), 158.5 (C15 or C19), 136.3 (C5), 136.2 (C5), 131.6 (C6), 126.7 (C7 or C8), 126.5 (C7 or C8), 117.5 (C7 or C8), 116.4 (C7 or C8), 104.5 (C13 or C16), 104.3 (C13 or C16), 98.8 (C13 or C16), 98.4, 60.5 (C1), 59.3 (C17 or C18), 59.3 (C17 or C18), 55.6 (C11 or C10), 55.5 (C17 or C18), 55.4, 49.3, 45.3, 35.8 (C2 or C4), 33.5 (C2 or C4), 31.1, 27.4 (C2 or C4), 27.1, 25.4, 25.3, 25.1, 24.1 (C3), 21.2 (C3).

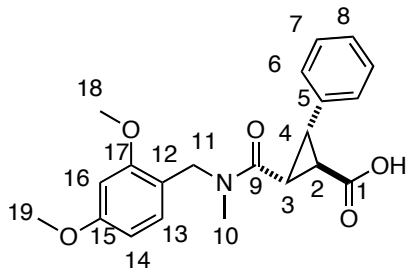


**(1*R*,2*R*,3*R*)-*N*-(2,4-dimethoxybenzyl)-2-formyl-*N*-methyl-3**

**phenylcyclopropanecarboxamide (2.26).** Tetrapropylammonium-perruthenate (TPAP) (5.1 mg, 0.014 mmol) was added to a suspension of *N*-methylmorpholine-*N*-oxide (128 mg, 1.10 mmol), 2.19 (312 mg, 0.730 mmol), and powdered, activated 4 Å-molecular sieves (730 mg) in CH<sub>2</sub>Cl<sub>2</sub> (7.3 mL). The reaction mixture was stirred at room temperature for 8 h. The black suspension was then filtered through a plug of silica gel (3cm x 5cm), and the plug was washed with CH<sub>2</sub>Cl<sub>2</sub>, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated under reduced pressure. The crude material was purified by flash chromatography eluting with EtOAc:Hex (1:1) to yield 216 mg (70%) of the titled compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) (rotamers 6:4) δ 9.48 (d, *J* = 7.5 Hz, 0.4 H), 9.44 (d, *J* = 7.5 Hz, 0.4 H), 7.23-7.07 (comp, 5 H), 6.95 (d, *J* = 8.5 Hz, 0.5 H), 6.39 (d, *J* = 8.5 Hz, 1.25 H), 6.35 (dt, *J* = 8.5, 2.1 Hz, 0.5 H), 4.95 (d, *J* = 16.4 Hz, 0.5 H), 4.58 (d, *J* = 16.4 Hz, 0.5 H), 4.39 (d, *J* = 11.6 Hz, 0.4 H), 4.37 (d, *J* = 11.6 Hz, 0.4 H), 4.35 (d, *J* = 11.6 Hz, 0.4 H), 3.71 (s, 6 H), 3.11 (s, 1.1 H), 2.97 (q, *J* = 8.6 Hz, 1 H), 2.87 (s, 1.4 H), 2.82-2.77 (comp, 0.6 H), 2.67 (t, *J* = 8.6 Hz, 0.4 H), 2.13-2.05 (m, 0.9 H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 177.0, 168.48, 167.1, 166.7, 160.7, 159.9, 158.3, 158.2, 134.9, 134.5, 133.2, 129.5, 129.0, 128.7, 128.6, 128.4, 128.2, 127.9, 127.1, 116.9, 116.2,

104.4, 103.9, 98.8, 98.0, 55.4, 48.4, 45.0, 34.8, 34.1, 33.33, 33.1 32.9, 32.7, 32.4, 32.2, 30.9, 30.8, 25.8, 25.1, 21.5; IR (neat) 2936 (br), 1723, 1696, 1611 cm<sup>-1</sup> ; mass spectrum (CI) *m/z* 354.1706 [C<sub>21</sub>H<sub>24</sub>NO<sub>4</sub> (M+1) requires 354.1705].

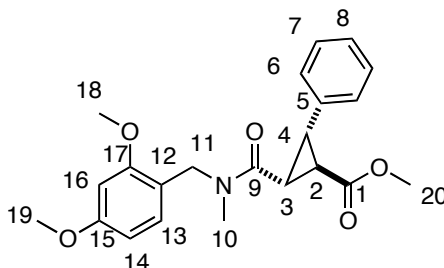
**NMR Assignments:** Rotamers (60/40) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.48 (d, *J* = 7.5 Hz, 0.4 H, C1-H), 9.44 (d, *J* = 7.5 Hz, 0.4 H, C1-H), 7.23-7.07 (comp, 5 H, C6-H & C7-H & C8-H & C15-H), 6.95 (d, *J* = 8.5 Hz, 0.5 H, C7-H), 6.35 (dt, *J* = 8.5, 2.1 Hz, 0.5 H, C14-H), 4.95 (d, *J* = 16.4 Hz, 0.5 H, C11-H), 4.58 (d, *J* = 16.4 Hz, 0.4 H, C11-H), 4.37 (d, *J* = 11.6 Hz, 0.4 H, C11-H), 4.35 (d, *J* = 11.6 Hz, 0.4 H), 3.71 (s, 6 H, C17-H & C18-H), 3.11 (s, 1.1 H, C10 H), 2.97 (m, 1H, C3-H), 2.87 (s, 1.4 H, C10-H), 2.82-2.77 (comp, 0.6 H, C2-H), 2.67 (t, *J* = 8.89 Hz, 0.4 H, C2-H), 2.13-2.05 (m, 1 H, C4-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 177.0 (C1), 168.5 (C9), 167.1, 166.7, 160.7, 159.9 (C16, C17, or C15), 158.3 (C16, C17, or C15), 158.2 (C16, C17, or C15), 134.9 (C5), 134.5, 133.2 (C6, C7, or C8), 129.5 (C6, C7, or C8), 129.0 (C6, C7, or C8), 128.7 (C6, C7, or C8), 128.6, 128.4 (C6, C7, or C8), 128.2, 127.9, 127.1, 116.9 (C18 or C19), 116.2 (C11), 104.4, 103.9, 98.8, 98.0, 55.4 (C18 or C19), 48.4 (C18 or C19), 45.0 (C11), 34.8 ( C4, or C10), 34.1 (C4, C10), 33.33 (C10 or C4), 33.16, 32.9 (C10 or C4), 32.7, 32.4, 32.2 (C2, C3, or C4), 30.94 (C2, C3, or C4), 30.79 (C2, C3, or C4), 25.8 (C2, C3, or C4), 25.1, 21.5



**[1-R-(1 $\alpha$ -2 $\alpha$ -3 $\alpha$ )]-2-(4-Phenyl)-3-[-(2,4-dimethoxy-benzyl)-methyl-carbamoyl]-cyclopropanecarboxylic acid (2.19).** A solution of the aldehyde **2.20** in MeOH (5.1 mL) containing Et<sub>3</sub>N (1.31 g, 10.2 mmol) was heated under reflux for 20 h. The solution was then cooled to room temperature whereupon the solvents were removed under reduced pressure. The residue was then dissolved in acetone (9.2 mL), whereupon 8 N Jones Reagent (230  $\mu$ L, 1.84 mmol) was added dropwise. After stirring the mixture for 45 min, a saturated aqueous solution of NaCl (5 mL) was added slowly. The mixture was extracted with EtOAc (5 mL), dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated under reduced pressure. The crude material was then dissolved in EtOAc (5 mL) and passed through a filter frit, whereupon the mother liqueur was concentrated under reduced pressure. The crude material was purified by silica gel chromatography eluting with EtOAc (1% AcOH) to yield 197 mg (97%) of the title product as an amber-colored oil. <sup>1</sup>H NMR (400 MHz CDCl<sub>3</sub>) (rotamers 6:4)  $\delta$  7.32-7.22 (comp, 8 H), 7.21-7.18 (comp, 6H) 6.82 (d, *J* = 8.5 Hz, 0.5 H), 6.50-6.44 (comp, 1.5 H), 6.40 (d, *J* = 2.4 Hz, 0.25 H), 6.38 (d, *J* = 2.4 Hz, 0.25 H), 6.35 (d, *J* = 2.4 Hz, 0.50 H), 6.16 (dd, *J* = 7.7, 2.4 Hz, 0.5 H), 4.91 (d, *J* = 15.0 Hz, 0.5 H), 4.54 (d, *J* = 15.0 Hz, 0.5 H) 4.21 (d, *J* = 15.0 Hz, 0.5 H), 4.03 (d, *J* = 15.0 Hz, 0.5), 3.80 (s), 3.76 (s), 3.71 (s), 3.15-2.95 (comp, 5.5 H), 2.86-2.78 (m, 1 H), 2.71 (s, 1.5 H), 2.35 (s, 2.9H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  177.0, 176.8, 168.5, 167.1, 166.7, 160.7, 159.9, 158.3, 158.250, 134.9, 134.5, 133.3, 129.5, 129.4,

129.0, 128.7, 128.6, 128.4, 128.3, 128.2, 127.9, 127.5, 127.1, 125.3, 116.9, 116.2, 104.4, 104.0, 98.8, 98.6, 98.3, 98.0, 55.4, 55.3, 55.2, 48.4, 45.1, 34.8, 34.1, 33.3, 33.2, 32.9, 32.7, 32.4, 32.2, 30.9, 30.7, 25.8, 25.1, 21.5. mass spectrum (CI)  $m/z$  370.1651 [ $C_{21}H_{24}NO_5$  (M+1) requires 370.1654].

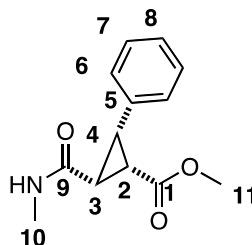
**NMR Assignments:**  $^1H$  NMR (400 MHz,  $CDCl_3$ ) (rotamers 60:40)  $\delta$  7.32-7.22 (comp, 8 H, C6-H & C8-H), 7.21-7.18 (comp, 6H, C7-H), 6.82 (d,  $J = 8.5$  Hz, 0.5H, C13-H), 6.50-6.44 (comp, 1.5 H, C16-H), 6.25 (d,  $J = 8.5$  Hz, C14-H), 6.16 (dd,  $J = 7.7$ , 2.4 Hz, C13-H, 0.5 H), 2.71 (s, 1.5-H, C10-H);  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  177.0 (C1 or C9), 176.8 (C1 or C9), 168.5 (C17 or C15), 167.1 (C17 or C15), 166.7 (C17 or C15), 160.7 (C17 or C15), 159.9 (C17 or C15), 158.3, 158.25 (C16 or C13), 134.9 (C5 or C8), 134.5 (C5 or C8), 133.3(C5 or C8), 129.5 (C6 or C7), 129.4 (C6 or C7), 129.0 (C6 or C7), 128.7 (C6 or C7), 128.6, 128.4 (C6 or C7), 128.3, 128.2, 127.9 (C19 or C13), 127.5 (C19 or C13) 127.1, 125.3 (C19 or C13), 116.9, 116.2 (C19 or C13), 104.4, 104.0, 98.8, 98.6, 98.3, 98.0, 55.4 (C18 or C17), 55.3 (C18 or C17), 55.2 (C18 or C17), 48.4 (C11 or C10), 45.1 (C11 or C10), 34.8 (C11 or C10), 34.1, (C11 or C10), 33.3 (C2, C3, C4), 33.2, 32.9 (C2, C3, C4), 32.7, 32.4 (C2, C3, C4), 32.2, 30.9, (C2, C3, C4) 30.7, 25.8, 25.1, 21.5.



**[1-R-(1 $\alpha$ -2 $\alpha$ -3 $\alpha$ )]-2-(4-Phenyl)-3-(2,4-dimethoxy-benzyl)-methyl-ester]-cyclopropanecarboxylic acid (**2.20**). SOCl<sub>2</sub> (12.6 mg, 0.293 mmol) was added to a solution of **2.21** (90.0 mg, 0.244 mmol) in MeOH (1.0 mL). The reaction mixture was stirred at 0 °C for 30 min and then at room temperature for 6 h. The solution was concentrated under reduced pressure, and the crude material was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub> (1% MeOH) to yield 32 mg (34%) of the title product as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) (rotamers 6:4)  $\delta$  7.32-7.22 (comp, 8 H), 7.21-7.18 (comp, 6 H) 6.76 (d,  $J$  = 8.5 Hz, 0.5 H), 6.44-6.42 (comp, 1.5 H), 6.37-6.31 (comp, 2.5 H), 6.17 (dd,  $J$  = 7.7 Hz, 2.4 Hz, 2 H), 4.93 (d,  $J$  = 15.4 Hz, 1 H), 4.51 (d,  $J$  = 15.4 Hz, 1 H) 4.21 (d,  $J$  = 15.4 Hz, 1 H), 4.03 (d,  $J$  = 16.1 Hz, 1H), 3.80 (s, 7 H), 3.76 (s, 7 H), 3.71 (s), 3.15-2.95 (comp, 4 H), 2.86-2.78 (comp, 6 H), 2.71 (s, 1.5 H), 2.35 (s, 3 H); <sup>13</sup>C (100 MHz, CDCl<sub>3</sub>)  $\delta$  183.1, 173.5, 166.5, 160.8, 160.0, 158.5, 155.6, 153.1, 143.0, 135.3, 134.9, 129.5, 128.9, 128.9, 128.6, 128.5, 128.2, 127.3, 127.2, 104.6, 104.1, 55.6, 52.4, 48.4, 45.1, 35.0, 33.5, 33.1, 33.1, 32.9, 32.7, 32.2, 29.9, 26.0, 25.2. IR (CDCl<sub>3</sub>) 2951, 2360, 1729, 1644, 1613, 1507, 1456 cm<sup>-1</sup>. mass spectrum (CI)  $m/z$  384.1812 [C<sub>22</sub>H<sub>26</sub>NO<sub>5</sub> (M+1) requires 384.1811].**

**NMR Assignment.** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) (rotamers 6:4)  $\delta$  7.32-7.22 (comp, 8 H), 7.21-7.18 (comp, 6H) 6.76 (d,  $J$  = 8.5 Hz, 0.5H), 6.44-6.42 (comp, 1.5 H),

6.37-6.31 (comp, 2.5 H), 6.17 (dd,  $J = 7.7, 2.4$  Hz, 2 H), 4.93 (d,  $J = 15.4$  Hz, 1 H), 4.51 (d,  $J = 15.4$  Hz, 1 H) 4.21 (d,  $J = 15.4$  Hz, 1 H), 4.03 (d,  $J = 16.1$  Hz, 1H), 3.80 (s, 7 H), 3.76 (s, 7 H), 3.71 (s), 3.15-2.95 (m, 4 H), 2.86-2.78 (m, 6.5 H), 2.71 (s, 1.6 H), 2.35 (s, 2.9H)  $^{13}\text{C}$  (100 MHz,  $\text{CDCl}_3$ )  $\delta$  183.1 (C1 or C9), 173.5 (C1 or C9), 166.5 (C17 or C15), 160.8 (C17 or C15), 160.0 (C17 or C15), 158.5 (C16 or C14), 155.6 (C16 or C14), 153.1 (C16 or C14), 143.0 (C16 or C14), 135.3, 134.9, 129.5 (C8, C7, C6 or C5), 128.9 (C8, C7, C6 or C5), 128.9 (C8, C7, C6 or C5), 128.6 (C8, C7, C6 or C5), 128.5 (C8, C7, C6 or C5), 128.2 (C8, C7, C6 or C5), 127.3 (C8, C7, C6 or C5), 127.2 (C8, C7, C6 or C5), 104.6 (C12), 104.1 (C12), 55.6 (C18 or C19), 52.4 (C8, C7, C6 or C5) 48.4 (C10 or C11), 45.1 (C10 or C11), 35.0 (C2, C3, C4), 33.5 (C2, C3, C4), 33.1 (C2, C3, or C4), 33.1, 32.9, 32.7, 32.2, 29.9(C3, C2 or C4) , 26.0, 25.2.

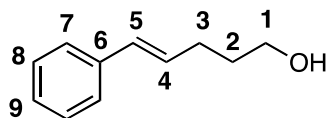


**[1-R-(1 $\alpha$ -2 $\alpha$ -3 $\alpha$ )]-2-(4-Phenyl)-3-hydroxymethylcyclopropanecarboxylic acid-(2,4-dimethoxybenzyl)methylamide (2.8).** A solution of the methyl ester (0.084 M) in  $\text{CH}_2\text{Cl}_2$  (835  $\mu\text{L}$ ) containing TFA (590 mg) was stirred at room temperature for 3h, whereupon the solution was concentrated under reduced pressure. The crude product was diluted with MeOH and filtered through a frit, and then concentrated under reduced pressure to yield 19 mg (100%) of the tilted product as a clear oil.  $^1\text{H}$  NMR (400 MHz,



CDCl<sub>3</sub>)  $\delta$  7.26-7.18 (comp, 5 H), 3.75 (s, 3 H), 2.92 (d,  $J$  = 5.5 Hz, 1 H), 2.80 (t,  $J$  = 5.5 Hz, 1 H), 2.57 (s, 3 H), 2.55-2.51 (m, 1 H); <sup>13</sup>C NMR (150 MHz, CD<sub>3</sub>OD),  $\delta$  174.2, 169.8, 136.2, 129.8, 129.2, 128.0, 52.8, 33.1, 32.4, 26.5, 25.6. mass spectrum (ESI)  $m/z$  278.1676 [C<sub>13</sub>H<sub>15</sub>NO<sub>3</sub> (M+1) requires 278.1673].

**NMR Assignments:** <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  7.26-7.18 (comp, 5 H, C6-H & C7-H & C8-H), 3.75 (s, 3H, C11-H), 2.92 (d,  $J$  = 5.5 Hz, 1 H, C4-H), 2.80 (t,  $J$  = 5.5 Hz, 1 H, C2-H), 2.57 (s, 3 H, C10-H), 2.55-2.51 (m, 1 H, C3-H); <sup>13</sup>C (150 MHz, CD<sub>3</sub>OD)  $\delta$  174.2 (C1), 169.8 (C9), 136.2 (C5), 129.8 (C7), 129.2 (C6), 128.0 (C8), 52.8 (C11), 33.1 (C4), 32.4 (C2), 26.5 (C3), 25.6 (C3).



**(E)-5-Phenylpent-4-en-1-ol (2.48).** A solution of **2.46** (0.18 g, 0.75 mmol) in THF (4 mL) was slowly added to a solution of LAH (0.029 g, 0.75 mmol) in THF (4 mL) at 0 °C. After gas evolution had subsided, the cooling bath was removed, and stirring was continued for 2 h. The mixture was recooled to 0 °C, whereupon sat. Na<sub>2</sub>SO<sub>4</sub> was slowly added until the formation of a white precipitant had subsided. The mixture was then filtered through a frit funnel, eluting with Et<sub>2</sub>O. The combined filtrate and washings were concentrated under reduced pressure. The crude residue was then purified by flash chromatography, eluting with CH<sub>2</sub>Cl<sub>2</sub> to afford 46 mg (35%) the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.39-7.19 (comp, 5 H), 6.42 (d,  $J$  = 15.7 Hz, 1 H),

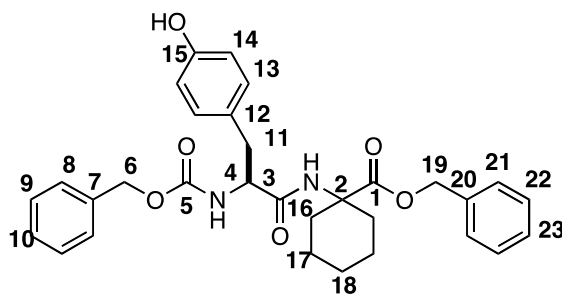
6.27-6.25 (comp, 1 H), 3.67-3.64 (comp, 2 H), 2.77-2.28 (comp, 2 H), 1.79-1.72 (comp, 2 H), 1.23-1.21 (m, 1 H);  $^{13}\text{C}$  (150 MHz) 137.6, 130.4, 130.0, 128.5, 126.9, 125.9, 62.4, 32.2, 29.3; IR (neat):, 3331.5 (br), 2918.9, 2359.3,  $\text{cm}^{-1}$ ; mass spectrum (ESI)  $m/z$  162.1044 [ $\text{C}_{16}\text{H}_{15}\text{O}_2$  requires 162.1045].

**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.39-7.19 (comp, 4 H, C7-H, C8-H, C9-H), 6.42 (d,  $J = 15.7$  Hz, 1 H, C5-H), 6.27-6.25 (comp, 1H, C4-H), 3.67-3.64 (m, 2 H, C1-H), 2.77-2.28 (m, 2 H, C3-H), 1.79-1.72 (comp, 2 H, C2-H), 1.23-1.21 (m, 1 H);  $^{13}\text{C}$  (150 MHz) 137.6 (C6), 130.4 (C8), 130.0 (C9), 128.5 (C7), 126.9 (C5), 125.9 (C4), 62.4 (C1), 32.2 (C2), 29.3 (C3).

#### **General procedures for the coupling of amino acids and peptides.**

**Method A:** *N*-Methylmorpholine (NMM) (602 mg, 5.04 mmol) was added to a solution of *N*-protected amino acid and *C*-protected amino acid (5.4 mmol) in DMF (17 mL) and stirred at  $-10^\circ\text{C}$ . The reaction mixture was stirred for 10 min at  $0^\circ\text{C}$ , whereupon 1-hydroxybenzotriazole hydrate (HOBt) (514 mg, 3.36 mmol) and 1-(3-Dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDCI) (355 mg, 1.85 mmol) were added. The ice bath was removed, and stirring was maintained for 14 h at room temperature. The mixture was concentrated to dryness under reduced pressure. Saturated aq.  $\text{NaHCO}_3$  (10 mL) was added, and the solid was isolated by filtration and washed with 1.0 *M* HCl (3 x 15 mL) and  $\text{H}_2\text{O}$  (15 mL). If the crude product was found not to be > 95% pure by  $^1\text{H}$  NMR, the crude product was purified by flash chromatography, recrystallization, or RP-HPLC using a binary grade of solvents, A and B (given).

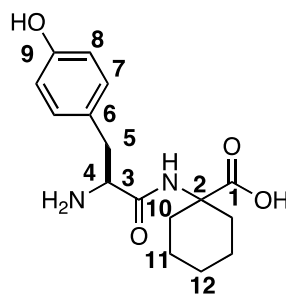
**Method B:** *N*-Methylmorpholine (NMM) (602 mg, 5.04 mmol) was added to a solution of *N*-protected amino acid and *C*-protected amino acid (5.4 mmol) in DMF (17 mL) and stirred at  $-10\text{ }^{\circ}\text{C}$ . The reaction mixture was stirred for 10 min at  $0\text{ }^{\circ}\text{C}$ , whereupon HOBt (514 mg, 3.36 mmol) and EDCI (355 mg, 1.85 mmol) were added. The ice bath was removed, and stirring was maintained for 14 h at room temperature. The mixture was concentrated to dryness under reduced pressure. Saturated aq.  $\text{NaHCO}_3$  (10 mL) was added and aqueous phase was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with 1.0 *M* HCl (3 x 15 mL) and  $\text{H}_2\text{O}$  (15 mL), dried ( $\text{Na}_2\text{SO}_4$ ), and concentrated to dryness under reduced pressure. If the crude product was found not to be > 95% pure by  $^1\text{H}$  NMR, the crude product was purified using flash chromatography, recrystallization, or RP-HPLC using a binary grade of solvents, A and B (given).



**(S)-Phenethyl-1-(2-(((benzyloxy)carbonyl)amino)-3-(4hydroxyphenyl)propanamido) cyclohexanecarboxylate (3.19).** Prepared from **3.18** and Cbz-Asn according to the general method B. The crude product was purified by silica gel chromatography eluting with  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (99:1) to yield 570 mg (64%) of the title compound as a clear oil.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.38-7.27 (comp, 10 H), 7.00 (d,  $J = 7.6$  Hz, 2 H), 6.70 (d,  $J = 7.6$  Hz, 2 H), 6.20 (br s, 1 H), 6.05 (br s, 1H), 5.41 (br s, 1 H), 5.16-5.03 (comp, 4 H), 4.33 (d,  $J = 6.4$  Hz, 1 H), 2.95 (dd,  $J = 13.6, 5.6$  Hz, 1 H),

2.85 (dd,  $J = 13.6, 5.6$  Hz, 1 H), 1.90 (d,  $J = 13.2$  Hz, 2 H), 1.76 (t,  $J = 11.6$  Hz, 2 H), 1.58-1.42 (comp, 3 H), 1.28-1.05 (comp, 3 H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  173.6, 170.2, 155.0, 135.8, 130.5, 128.5, 128.5, 128.2, 128.0, 115.6, 67.0, 58.9, 56.1, 37.2, 32.2, 31.9, 24.9, 21.2, 21.0.; mass spectrum (CI)  $m/z$  531.2490 [ $\text{C}_{31}\text{H}_{35}\text{N}_2\text{O}_6$  (M+1) requires 531.2495].

**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.38-7.27 (comp, 10 H, C8-H, C9-H, C10-H, C21-H, C22-H, C23-H), 7.00 (d,  $J = 7.6$  Hz, 2 H, C13-H), 6.70 (d,  $J = 7.6$  Hz, 2 H, C14-H), 5.41 (br s, 1 H, OH), 5.16-5.03 (comp, 4 H, C6-H and C19-H), 4.33 (d,  $J = 6.4$  Hz, 1 H, C4-H), 2.95 (dd,  $J = 13.6, 5.6$  Hz, 1 H, C11-H), 2.85 (dd,  $J = 13.6, 5.6$  Hz, 1 H, C11-H), 1.90 (d,  $J = 13.2$  Hz, 2 H, C16-H or C16'-H), 1.76 (app t,  $J = 11.6$  Hz, 2 H, C16-H or C16'-H), 1.58-1.42 (comp, 3 H, C17-H or C17'-H or C18-H), 1.28-1.05 (comp, 3 H, C17-H or C17'-H or C18-H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  173.6 (C3), 170.2 (C1), 155.0 (C5), 135.8 (C15), 130.5 (C7 & C24), 128.5 (C13 or C12 or C9 or C22), 128.5 (C13 or C12 or C9 or C22), 128.2 (C9 or C10 or C22 or C15), 128.0 (C8 and C21), 115.6 (C14), 67.1 (C6 or C19), 67.0 (C6 or C19), 58.9 (C3), 56.1 (C2), 37.2 (C16), 32.2 (C16 or C16'), 31.9 (C16 or C16'), 24.9 (C18), 21.2 (C17 or C17'), 21.0 (C17 or C17').

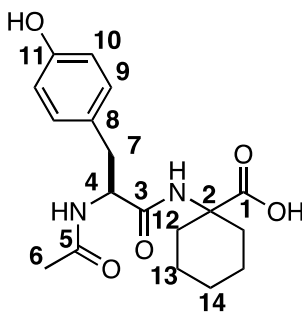


**(S)-1-(2-Amino-3-(4-hydroxyphenyl)propanamido)cyclohexanecarboxylic**

**acid. (3.20).** The benzyl carbamate **3.19** was dissolved in MeOH (1.5 mL) containing 10% Pd/C (10 mg, 10 wt %). The resulting mixture was purged four times with H<sub>2</sub>, and the suspension stirred under H<sub>2</sub> (1 atm) for 14 h. The mixture was filtered through a pad of celite, and the pad was washed with MeOH. The combined filtrate and washings were concentrated to dryness under reduced pressure to yield 59 mg (99%) of the title compound as a clear glass. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 7.13 (d, *J* = 8.4 Hz, 2 H), 6.76 (d, *J* = 8.4 Hz, 2 H), 4.04 (t, *J* = 7.6 Hz, 1 H), 3.05 (dd, *J* = 13.8, 8.2 Hz, 1 H), 2.96 (dd, *J* = 13.8, 8.2 Hz, 1 H), 2.23 (d, *J* = 13.6 Hz, 1 H), 1.85-1.76 (comp, 2 H), 1.67-1.63 (m, 1 H), 1.52-1.12 (comp, 5 H), 1.28-1.05 (q, *J* = 12.8 Hz, 1 H); <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD): δ 181.3, 169.6, 158.0, 131.6, 126.9, 116.7, 62.9, 56.2, 38.1, 35.4, 31.7, 26.6, 22.8, 22.7. Mass spectrum (ESI) *m/z* 307.1652 [C<sub>16</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub> (M+H)<sup>+</sup> requires 307.1657].

**NMR Assignment.** <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.13 (d, *J* = 8.4 Hz, 2H, C7-H), 6.76 (d, *J* = 8.4 Hz, 2H, C8-H), 4.04 (t, *J* = 7.6 Hz, 1 H, C4-H), 3.05 (dd, *J* = 13.8, 8.2 Hz, 1H, C5-H), 2.96 (dd, *J* = 13.8, 8.2 Hz, 1H, C5-H), 2.23 (app d, *J* = 13.6 Hz, 1H, C10-H or C10'-H), 1.85-1.76 (comp, 2 H, C10-H or C10'-H), 1.66 (m, 1H, C10-H or C10'-H), 1.52-1.12 (comp, 5 H, C11-H and C12-H and C11'-H), 1.28-1.05 (q, *J* = 12.8 Hz, 1 H, C11-H or C12-H or C11'-H); <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 181.3 (C1), 169.6 (C3), 158.0 (C9), 131.6 (C7), 126.9 (C6), 116.7 (C8), 62.9 (C2), 56.2 (C3), 38.1 (C5), 35.4

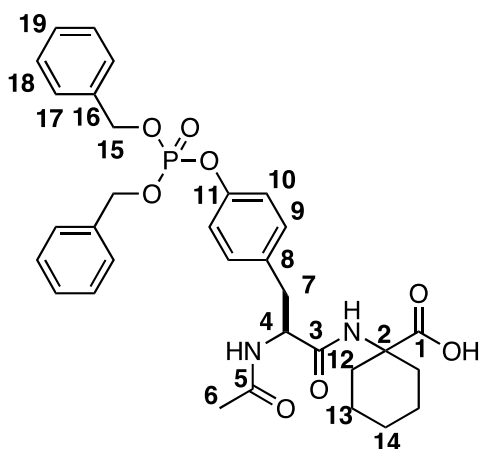
(C10 or C10'), 31.7 (C10 or C10'), 26.6 (C11, C12, or C11'), 22.8 (C11, C12 or C11'), 22.7 (C11 or C12 or C11').



**(S)-1-(2-Acetamido-3-(4-(phosphonoxy)phenyl)propanamido)**

**cyclohexanecarboxylic acid. (3.21)** Acetic anhydride was added to a solution of **3.20** (634 mg, 2.07 mmol) in dioxane/H<sub>2</sub>O (1:1, 60 mL). This mixture was stirred at room temperature for 30 min when it was then concentrated to dryness under reduced pressure. Azeotropic removal of Ac<sub>2</sub>O was accomplished via the addition of toluene (1.5 mL) to the reaction vessel followed by concentration under reduced pressure. This procedure was repeated two more times followed by drying *in vacuo*, which afforded 743 mg (100 %) of product as a glass. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.08 (d, *J* = 8.4 Hz, 2 H), 6.79 (d, *J* = 8.4 Hz, 2 H), 4.64 (dd, *J* = 8.6, 6.2 Hz, 1 H), 3.00 (dd, *J* = 13.8, 6.2 Hz, 1 H), 2.76 (dd, *J* = 13.8, 6.2 Hz, 1 H), 2.08-1.97 (comp, 2 H), 1.90 (s, 3 H), 1.76 (td, *J* = 12.8, 3.6 Hz, 2 H), 1.61-1.21 (comp, 6 H); <sup>13</sup>C NMR (100 MHz, CD<sub>3</sub>OD) δ 177.8, 173.2, 173.0, 157.2, 131.4, 129.2, 116.2, 60.3, 56.0, 38.2, 33.4, 33.0, 26.4, 22.5, 22.4. Mass spectrum (ESI) *m/z* 349.1758 [C<sub>18</sub>H<sub>25</sub>N<sub>2</sub>O<sub>5</sub> (M+H) requires 349.1763].

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.08 (d,  $J = 8.4$  Hz, 2 H, C9-H), 6.79 (d,  $J = 8.4$  Hz, 2 H, C10-H), 4.64 (dd,  $J = 8.6, 6.2$  Hz, 1 H, C4-H), 3.00 (dd,  $J = 13.8, 6.2$  Hz, 1 H, C7-H), 2.76 (dd,  $J = 13.8, 6.2$  Hz, 1 H, C7-H), 2.08-1.97 (comp, 2 H, C12-H or C12'-H), 1.90 (s, 3 H, C6-H), 1.76 (td,  $J = 12.8, 3.6$  Hz, 2 H, C12-H or C12'-H), 1.61-1.21 (comp, 6 H, C13-H, C14-H, and C13'-H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  177.8 (C1), 173.2 (C3), 173.0 (C5), 157.2 (C11), 131.4 (C9), 129.2 (C8), 116.2 (C10), 60.3 (C2), 56.0 (C4), 38.2 (C7), 33.4 (C12, C12' or C14 or C6), 33.0 (C12, C12' or C14 or C6), 26.4 (C12, C12' or C14 or C6), 22.5 (C13 & C13'), 22.4 (C12, C12' or C14 or C6).



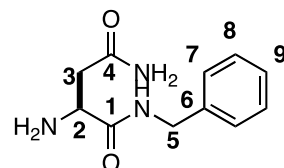
**(S)-1-(2-Acetamido-3-(4-(phosphonooxy)phenyl)propanamido) cyclohexane-carboxylic acid (3.15).** 1-H-Tetrazole (785 mg, 11.2 mmol) and diisopropylphosphoramidite were added to a solution of **3.21** in DMF (21 mL) at 0 °C, and the solution was stirred at 0 °C for 1 h and then at room temperature for 15 h. The solution was cooled to 0 °C, and 6 M *tert*-butyl hydroperoxide in decane (5.3 mL) was added. The resulting solution was stirred at 0 °C for 30 min and then at room temperature for 5 h, whereupon it was cooled to 0 °C and 5% aqueous  $\text{NaHSO}_3$  (5 mL) was added. The solution was stirred at 0 °C for 30 min and then at room temperature for 2 h. The

mixture was transferred to a separatory funnel containing H<sub>2</sub>O (20 mL), and the layers were separated. The aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (15 mL x 3) whereupon the combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure. The residue was triturated with Et<sub>2</sub>O (4 x 5 mL) to yield 250 mg (20%) of the title compound. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.69 (br s, 1 H), 7.34-7.27 (comp, 10 H), 7.18 (d, *J* = 8.4 Hz, 2 H), 7.06 (d, *J* = 8.4 Hz, 2 H), 5.08 (app d, *J* = 6.4 Hz, 4 H), 4.83 (dd, *J* = 14.4, 6.8 Hz, 1 H), 3.11-3.97 (comp, 2 H), 1.98-1.90 (comp, 2H), 1.93 (s, 3 H), 1.77-1.65 (comp, 2 H), 1.56-1.02 (comp, 6 H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ 176.7, 171.5, 171.4, 149.3, 149.2, 135.3, 135.2, 133.8, 130.8, 128.7, 128.6, 128.0, 120.0, 119.98, 70.1, 70.0, 59.1, 54.4, 36.8, 32.4, 31.4, 25.0, 22.7, 21.2, 21.1. Mass spectrum (ESI) *m/z* 607.2211 [C<sub>32</sub>H<sub>36</sub>N<sub>2</sub>O<sub>8</sub>P (M+H) requires 607.2215].

**NMR Assignments.** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.69 (br s, 1H, N-H), 7.34-7.27 (comp, 10 H, C17-H, C18-H, C19-H), 7.18 (d, *J* = 8.4 Hz, 2 H, C9-H), 7.06 (d, *J* = 8.4 Hz, 2 H, C10-H), 5.08 (app d, *J* = 6.4 Hz, 4 H, C15-H), 4.83 (dd, *J* = 14.4, 6.8 Hz, 1 H, C4-H), 3.10-2.97 (comp, 2 H, C7-H), 1.98-1.90 (comp, 2 H, C12'-H or C12-H), 1.93 (s, 3 H, C6-H), 1.77-1.65 (comp, 2 H, C12-H or C12'-H), 1.56-1.02 (comp, 6 H, C13-H & C14-H & C13'-H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ 176.7 (C1), 171.5 (C5 or C3), 171.4 (C5 or C3), 149.3 (C11 or C8), 149.2 (C11 or C8), 135.3 (C15), 135.2 (C15), 133.8 (C9, C16, or C19), 130.8 (C9 or C16 or C19), 128.7 (C9 or C16 or C19), 128.6 (C18), 128.0 (C17), 120.0 (C10), 119.98 (C10), 70.1 (C3), 59.1 (C4), 54.4 (C2), 36.8 (C7), 32.4 (C12 or C12'), 31.4 (C12 or C12'), 25.0 (C6 or C13 or C14 or C13'), 22.7 (C6 or C13 or C14 or C13'), 21.2 (C6 or C13 or C14 or C13'), 21.1 (C6 or C13 or C14 or C13').

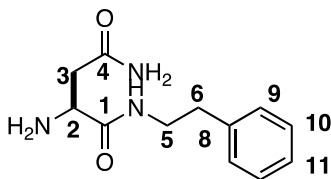


**Representative procedure for the synthesis of asparagine residues.**



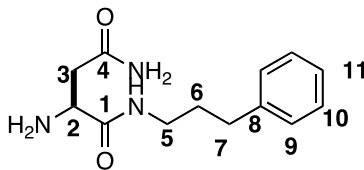
**2(S)-Amino-*N'*-benzylsuccinamide (3.43).** Prepared from **3.23** and Cbz-Asn according to the general method A. The crude product was dissolved in MeOH containing 10% Pd/C (18 mg), and the mixture was purged four times with H<sub>2</sub>. The suspension was stirred under H<sub>2</sub> (1 atm) for 2 h. The mixture was filtered through a pad of celite, and the pad was washed with MeOH (5.2 mL). The combined filtrate and washings were concentrated under reduced pressure to yield 78 mg (64%) of the title compound as a white solid over two steps: mp: 143-144 °C. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.31-7.29 (comp, 3 H), 7.25-7.21 (comp, 2 H), 3.70 (dd, *J* = 8.0, 5.1 Hz, 1 H), 3.31-3.30 (comp, 2 H), 2.61 (dd, *J* = 15.5, 8.0 Hz, 1 H), 2.47 (dd, *J* = 15.5, 8.0 Hz, 1 H), 1.82 (p, *J* = 7.5 Hz, 2 H); <sup>13</sup>C (150 MHz) 175.1, 174.7, 138.6, 128.4, 127.4, 127.0, 52.2, 42.9, 39.9. mass spectrum (ESI) *m/z* 222.1237 [C<sub>11</sub>H<sub>15</sub>N<sub>3</sub>O<sub>2</sub> (M+1) requires 222.1240].

**NMR Assignments.** <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.31-7.29 (comp, 3 H, C8-H and C9-H), 7.30-7.23 (comp, 2 H, C7-H), 3.70 (dd, *J* = 8.0, 5.1 Hz, C2-H), 3.31-3.30 (comp, 2 H, C5-H), 2.61 (dd, *J* = 15.4, 8.0 Hz, 1 H, C3-H), 2.47 (dd, *J* = 15.2, 8.0 Hz, 1 H, C3-H). <sup>13</sup>C (150 MHz) 175.1 (C1 or C2), 174.7 (C1 or C2), 138.6 (C6), 128.4 (C6), 127.4 (C7), 127.0 (C9), 52.2 (C2), 42.9 (C5), 39.9 (C3).



**2(S)-Amino-*N'*-phenethylsuccinamide (3.32).** Prepared from **3.24** according to the representative procedure for the synthesis of asparagine residues to yield 168 mg (94%) of the title compound as a white solid over two steps. mp = 143-144 °C. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.30-7.26 (comp, 3 H), 7.24-7.18 (comp, 2 H), 3.68 (dd, *J* = 8.0, 5.1 Hz, 1 H), 3.46-3.40 (comp, 2H), 2.61 (dd, *J* = 15.4, 8.0 Hz, 1 H), 2.47 (dd, *J* = 15.2, 8.0 Hz, 1 H); <sup>13</sup>C (150 MHz) 175.1, 174.5, 139.3, 128.7, 128.4, 126.2, 52.0, 40.8, 39.5, 35.3. mass spectrum (ESI) *m/z* 236.1394 [C<sub>12</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub> (M+1) requires 236.1394].

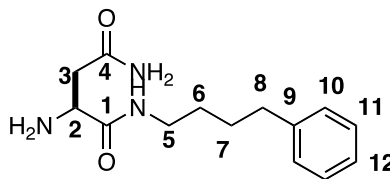
**NMR Assignments.** <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.30-7.26 (comp, 3 H, C10-H and C11-H), 7.24-7.18 (comp, 2 H, C9-H), 3.68 (dd, *J* = 8.0, 5.1 Hz C2-H), 3.46-3.40 (comp, 2 H, C5-H), 2.81 (t, *J* = 7.1 Hz, 2 H), 2.61 (dd, *J* = 15.4, 8.0 Hz, 1 H, C3-H), 2.47 (dd, *J* = 15.2, 8.0 Hz, 1 H, C3-H), 1.82 (p, *J* = 7.52 Hz, 2H, ); <sup>13</sup>C (150 MHz) 175.1 (C1 or C4), 174.5 (C1 or C4), 139.3 (C8), 128.7 (C10), 128.4 (C9), 126.2 (C11), 52.0 (C2), 40.8 (C5), 39.5 (C3), 35.3 (C6).



**2(S)-Amino-*N'*-(3-phenylpropyl)succinamide (3.33).** Prepared from **3.26** according to the representative procedure for the synthesis of asparagine residues to yield 18 mg (24%) of the title compound as a white solid over two steps. mp = 170-172 °C. <sup>1</sup>H

NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  7.27-7.14 (comp, 5 H), 3.64 (dd,  $J$  = 8.0, 5.5 Hz, 1 H), 3.23 (dd,  $J$  = 10.0, 5.5 Hz, 2 H), 2.66-2.58 (comp, 3 H), 2.46 (dd,  $J$  = 15.0, 8.0 Hz, 1 H), 1.82 (p,  $J$  = 7.52 Hz, 2 H); <sup>13</sup>C (150 MHz) 176.2, 175.8, 143.0, 129.5, 129.4, 126.9, 53.3, 41.2, 40.0, 34.1, 32.2. mass spectrum (ESI)  $m/z$  250.1304 [C<sub>13</sub>H<sub>17</sub>N<sub>4</sub>O<sub>2</sub> (M+1) requires 250.1304].

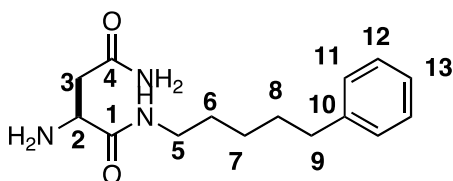
**NMR Assignments.** <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  7.27-7.14 (comp, 5 H, C9–H and C10–H and C11–H), 3.64 (dd,  $J$  = 8.0, 5.5 Hz, 1 H, C2-H), 3.23 (dd,  $J$  = 10.0, 5.5 Hz, 2 H, C5-H), 2.66-2.58 (comp, 3 H, C7-H and C3-H), 2.46 (dd,  $J$  = 15.0, 8.0 Hz, 1 H, C3-H) 1.82 (p,  $J$  = 7.52 Hz, 2 H, C6–H); <sup>13</sup>C (150 MHz) 176.2 (C1 or C4), 175.8 (C1 or C4), 143.0 (C8), 129.5 (C10), 129.4 (C9), 126.9 (C11), 53.3 (C2), 41.2 (C5), 40.0 (C3), 34.1 (C7), 32.2 (C6).



**2(S)-Amino-*N'*-(4-phenylbutyl)succinamide (3.34).** Prepared from **3.26** according to the representative procedure for the synthesis of asparagine residues to yield 20 mg (50%) of the title compound as a white solid over two steps: mp = 119-121 °C. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  7.26-7.13 (comp, 5 H), 3.62 (dd,  $J$  = 8.0, 5.5 Hz, 1 H), 3.23-3.19 (comp, 2 H), 2.66-2.58 (comp, 3 H), 2.46 (dd,  $J$  = 15.0, 8.0 Hz, 1 H), 1.82 (p,  $J$  = 7.5 Hz, 2 H), 1.56 (p,  $J$  = 7.5 Hz, 2 H); <sup>13</sup>C (150 MHz)  $\delta$  176.2, 175.8, 143.0, 129.5, 129.4,

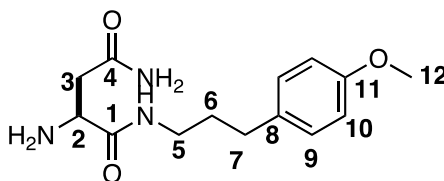
126.9, 53.3, 41.2, 40.0, 34.1, 32.2, 30.9. mass spectrum (ESI)  $m/z$  286.15265 [ $C_{14}H_{21}N_3O_2$  (M+Na) requires 286.1526].

**NMR Assignments.**  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta$  7.26-7.13 (comp, 5 H, C10-H, C11-H, and C12-H), 3.62 (dd,  $J = 8.0$  Hz, 5.5 Hz, 1 H, C2-H), 3.23-3.19 (comp, 2H, C5-H), 2.66-2.58 (comp, 3H, C8-H and C3-H), 2.46 (dd,  $J = 15.0$ , 8.0 Hz, 1H, C3-H), 1.82 (p,  $J = 7.5$  Hz, 2H, C6-H or C7-H), 1.56 (p,  $J = 7.5$  Hz, 2H, C6-H or C7-H);  $^{13}C$  (150 MHz)  $\delta$  176.2 (C1 or C4), 175.8 (C1 or C4), 143.0 (C10, C11 or C12), 129.5 (C10, C11 or C12), 129.4 (C10, C11 or C12), 126.9 (C10, C11 or C12), 53.3 (C2), 41.2 (C5), 40.0 (C8), 34.1 (C3), 32.2 (C6), 30.9 (C7).



**2(S)-Amino-*N'*-(5-phenylpentyl)succinamide (3.35)** Prepared from **3.27** according to the representative procedure for the synthesis of asparagine residues to yield 29 mg (70%) of the title compound as a glass over two steps.  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta$  7.11-7.10 (comp, 5 H), 3.63 (dd,  $J = 8.0$ , 5.2 Hz, 1 H), 3.22 (dd,  $J = 14.2$ , 6.8 Hz, 1 H), 3.15 (dd,  $J = 14.2$ , 7.2 Hz, 1 H), 2.63-2.57 (comp, 3 H), 2.43 (dd,  $J = 15.2$ , 8.0 Hz, 1 H), 1.63 (p,  $J = 7.6$ , 2 H), 1.54 (p,  $J = 7.2$  Hz, 2 H), 1.40-1.32 (comp, 2 H);  $^{13}C$  NMR (100 MHz,  $CD_3OD$ ):  $\delta$  176.3, 175.9, 143.8, 129.4 (2C), 129.3 (2C), 126.7, 53.4, 41.3, 40.4, 36.8, 32.4, 30.3, 27.5. Mass spectrum (ESI)  $m/z$  278.1863 [ $C_{15}H_{24}N_3O_2$  (M+H) $^+$  requires 278.1860].

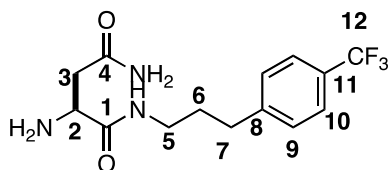
**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  7.11-7.10 (comp, 5 H, C11-H, and C12-H, and C13-H), 3.63 (dd,  $J$  = 8.0, 5.2 Hz, 1 H, C2-H), 3.22 (dd,  $J$  = 14.2, 5.2 Hz, 1 H, C5-H), 3.15 (dd,  $J$  = 14.2, 8.0 Hz, 1 H, C3-H), 2.63-2.57 (comp, 3 H, C3-H and C9-H), 2.43 (dd,  $J$  = 14.2, 8.0 Hz, 1 H), 1.63 (p,  $J$  = 7.6, 2 H, C6-H or C8-H), 1.54 (p,  $J$  = 7.6 Hz, 2 H, C6-H or C8-H), 1.40-1.32 (comp, 2 H, C7-H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  176.3 (C1 or C4), 175.9 (C1 or C4), 143.8 (C11), 129.4 (2C), 129.3 (2C), 126.7 (, 53.4 (C1), 41.3 (C5), 40.4 (C3), 36.8 (C9), 32.4 (C6, C7, or C8), 30.3 (C6, C7, or C8), 27.5 (C6, C7, or C8).



**2(S)-Amino-*N'*-(3-(4-methoxyphenyl)propyl)succinamide (3.69).** Prepared from **3.59** according to the representative procedure for the synthesis of asparagine residues to yield 83 mg (99%) of the title compound as a white solid: mp = 204–205 °C.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.10 (d,  $J$  = 6.5 Hz, 2 H), 6.8 (d,  $J$  = 6.5 Hz, 2 H), 3.66-3.62 (comp, 1 H), 3.23-3.19 (comp, 2 H), 2.66-2.44 (comp, 4 H), 2.15, (s, 3 H), 1.78 (p,  $J$  = 7.1 Hz, 2 H);  $^{13}\text{C}$  (150 MHz)  $\delta$  176.2, 175.8, 143.0, 129.5, 129.4, 126.9, 53.3, 41.2, 40.0, 34.1, 32.2, 30.9. mass spectrum (ESI)  $m/z$  280.1658 [ $\text{C}_{14}\text{H}_{22}\text{N}_3\text{O}_3$  (M+1) requires 280.1656].

**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.10 (d,  $J$  = 6.5 Hz, 2 H, C9-H), 6.8 (d,  $J$  = 6.5 Hz, 2 H, C10-H), 3.66-3.62 (comp, 1 H, C2-H), 3.23-3.19 (comp, 2 H,

C5-H), 2.66-2.44 (comp, 4 H, C7-H and C5-H), 2.15 (s, 3H, C12-H), 1.78 (p,  $J = 7.11$  Hz, 2 H, C6-H);  $^{13}\text{C}$  (150 MHz) 176.2 (C1-H or C4-H), 175.8 (C1-H or C4-H), 143.0 (C11), 129.5 (C9), 129.4 (C8), 126.9 (C10), 53.3 (C12), 41.2 (C5-H), 40.0 (C7-H), 34.1 (C3), 32.2 (C6), 30.9 (C6-H).

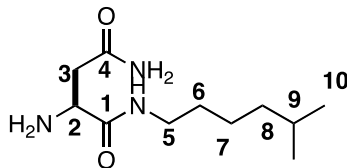


**2(S)-Amino-*N'*-(3-(4-(trifluoromethyl)phenyl)propyl)succinamide (3.68).**

Prepared from **3.58** according to the representative procedure for the synthesis of asparagine residues to yield 69 mg (65%) of the title compound as a white solid over two steps. mp = 119-120 °C.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.55 (d,  $J = 8.0$  Hz, 2 H), 7.40 (d,  $J = 8.0$  Hz, 2 H), 3.47 (dd,  $J = 7.1, 5.5$  Hz, 1 H), 3.24-3.19 (comp, 2 H), 2.72 (t,  $J = 7.8$  Hz, 2 H), 2.62 (dd,  $J = 15.4, 7.1$  Hz, 1 H), 2.62 (dd,  $J = 15.4, 7.1$  Hz, 1 H) 1.82 (p,  $J = 7.52$  Hz, 2 H);  $^{13}\text{C}$  (150 MHz)  $\delta$  210.0, 176.2, 175.8, 147.8, 130.1, 129.2, 127.0, 126.2, 124.8, 122.7, 41.0, 39.8, 33.8, 31.9, 30.7. mass spectrum (ESI)  $m/z$  318.1421 [ $\text{C}_{14}\text{H}_{19}\text{N}_3\text{O}_2\text{F}_3$  (M+1) requires 318.1421].

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.55 (d,  $J = 8.0$  Hz, 2 H, C9-H), 7.40 (d,  $J = 8.0$  Hz, 2 H, C10-H), 3.47 (dd,  $J = 7.1$  Hz, 5.5 Hz, 1 H, C2-H), 3.24-3.19 (comp, 2 H, C5-H), 2.72 (t,  $J = 7.8$  Hz, 2 H, C7-H), 2.64 (dd,  $J = 15.4, 7.1$  Hz, 1 H, C3-H), 2.62 (dd,  $J = 15.4, 7.1$  Hz, 1 H, C3-H) 1.82 (p,  $J = 7.52$  Hz, 2 H);  $^{13}\text{C}$  (150 MHz)  $\delta$  210.0 (C12), 176.2 (C1 and C4), 175.8 (C1 and C4), 147.8 (C11), 130.1 (C-Ar), 129.2

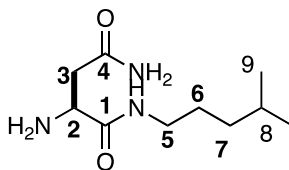
(C-Ar), 127.0 (C-Ar), 126.2 (C-Ar), 124.8 (C-Ar), 122.7 (C-Ar), 41.0 (C3), 39.8 (C5),  
33.8 (C3), 31.9 (C7), 30.7 (C6).



**2(S)-Amino-*N'*-(5-methylhexyl)succinamide (3.64).** Prepared from **3.54** according to the representative procedure for the synthesis of asparagine residues to yield 83 mg (>99%) of the title compound as a white solid over two steps: decomp at 210 °C.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  3.61 (dd,  $J = 8.2, 5.1$  Hz, 1 H), 3.24-3.12 (comp, 2 H), 2.62 (dd,  $J = 15.2, 8.2$  Hz, 1 H), 2.42 (dd,  $J = 15.2, 8.2$  Hz, 2 H), 1.59-1.47 (comp, 3 H), 1.24-1.20 (comp, 1-H), 0.89 (d,  $J = 6.5$  Hz);  $\delta$   $^{13}\text{C}$  (150 MHz) 221.5, 39.3, 39.1, 29.5, 27.9, 24.6, 21.8. mass spectrum (ESI)  $m/z$  252.1682 [ $\text{C}_{11}\text{H}_{23}\text{N}_3\text{O}_2$  (M+Na+1) requires 252.1683].

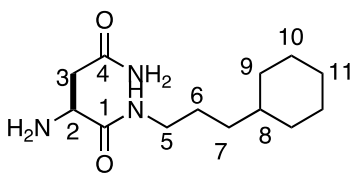
**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  3.61 (dd,  $J = 8.2, 5.1$  Hz, 1 H, C2-H), 3.24-3.12 (comp, 2 H, C5-H), 2.62 (dd,  $J = 15.2, 8.2$  Hz, 1 H, C3-H), 2.42 (dd,  $J = 15.2, 8.2$  Hz, 1 H, C3-H), 1.59-1.47 (comp, 3 H, C6-H or C7-H), 1.38-1.24 (comp, 2 H, C6-H or C7-H), 1.24-1.20 (comp, 3 H, C9-H and C8-H), 0.89 (d,  $J = 6.5$  Hz, 6 H, C10-H);  $\delta$   $^{13}\text{C}$  (150 MHz) 221.5 (C1 and C4), 39.3, (C5), 39.1 (C6), 39.1 (C3), 29.5 (C7, C8, C9, or C10), 27.9 (C7, C8, C9, or C10), 24.6 (C7, C8, C9, or C10), 21.8 (C7, C8, C9, or C10).





**2(S)-Amino-*N'*-(4-methylpentyl)succinamide (3.65).** Prepared from **3.55** according to the representative procedure for the synthesis of asparagine residues to yield 89 mg (67%) of the title compound as a glass over two steps.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  3.41 (dd,  $J = 7.1, 5.5$  Hz, 1 H), 3.24-3.12 (comp, 2H), 2.62 (dd,  $J = 15.2, 7.1$  Hz, 1 H), 2.53-2.19 (comp, 2 H), 1.59-1.47 (comp, 3 H), 1.24-1.20 (comp, 1 H), 0.89 (d,  $J = 6.5$  Hz). mass spectrum (ESI)  $m/z$  216.1706 [ $\text{C}_{10}\text{H}_{22}\text{N}_3\text{O}_2$  (M+1) requires 216.1705].

**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  3.41 (dd,  $J = 7.1, 5.5$  Hz, 1 H, C2-H), 3.24-3.12 (comp, 2 H, C5-H), 2.62 (dd,  $J = 15.2, 7.1$  Hz, 1 H, C3-H), 2.53-2.19 (comp, 2 H, C6-H), 1.59-1.47 (comp, 2 H, C7-H), 1.24-1.20 (comp, 1 H, C8-H), 0.89 (d,  $J = 6.5$  Hz, C9-H).

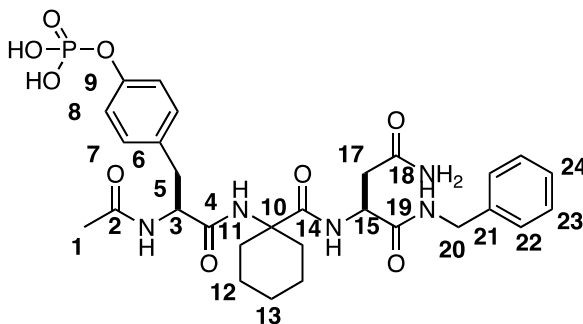


**2(S)-Amino-*N'*-(3-cyclohexylpropyl)succinamide (3.63).** Prepared from **3.53** according to the representative procedure for the synthesis of asparagine residues to yield 800 mg (72%) of the title compound as a white solid: mp = 167-168 °C.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  4.13 (dd,  $J = 8.5, 4.8$  Hz, 1 H), 3.25-3.12 (comp, 2 H), 2.84 (dd,  $J = 17.2, 4.8$  Hz, 1 H), 2.74 (dd,  $J = 17.2, 4.8$  Hz, 1 H), 1.79-1.62 (comp, 5 H), 1.55-1.51 (comp, 2 H), 1.27-1.18 (comp, 6 H), 0.94-0.84 (comp, 2 H);  $\delta$   $^{13}\text{C}$  (150 MHz) 221.5,

211.3, 39.8, 37.5, 35.1, 34.5, 33.3, 26.55, 26.45, 26.23. mass spectrum (ESI)  $m/z$  256.2025 [ $C_{13}H_{26}N_3O_2$  ( $M+Na+1$ ) requires 256.2020].

**NMR Assignment.**  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta$  4.13 (dd,  $J = 8.5, 4.8$  Hz, 1 H, C2-H), 3.25-3.12 (comp, 2 H, C5-H), 2.84 (dd,  $J = 17.2, 4.8$  Hz, 1 H, C3-H), 2.74 dd,  $J = 17.2, 4.8$  Hz, 1 H, C3-H), 1.79-1.62 (comp, 5 H, C6-H, C7-H, C8-H, C9-H, or C10-H), 1.55-1.51 (comp, 2 H, C6-H, C7-H, C8-H, C9-H, or C10-H), 1.27-1.18 (comp, 6 H, C6-H, C7-H, C8-H, C9-H, or C10-H), 0.94-0.84 (comp, 2 H, C6-H, C7-H, C8-H, C9-H, or C10-H);  $\delta$   $^{13}C$  (150 MHz) 221.5, (C1 or C4) 211.3 (C1 or C4), 39.8 (C5), 37.5 (C3), 35.1 (C6), 34.5 (C7, C8, C9, or C10), 33.3 (C7, C8, C9, or C10), 26.55 (C7, C8, C9, or C10), 26.45, (C7, C8, C9, or C10) 26.23 (C7, C8, C9, or C10).

#### Representative procedure for the synthesis of desired tripeptides.

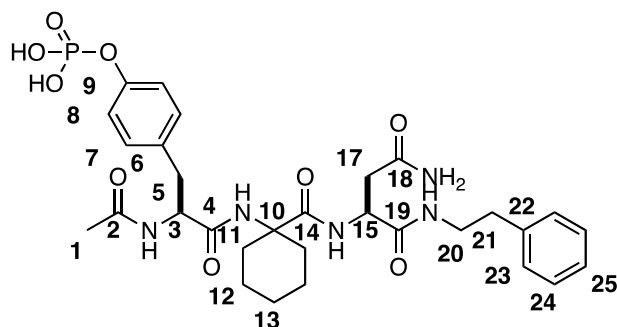


**4-((S)-2-Acetamido-3-((1-(((S)-4-amino-1-(benzyl amino)-1,4-dioxobutan-2-yl)cyclohexyl)amino)-3-oxopropyl)phenyl dehydrogenate phosphate (3.06).** Prepared from **3.31** and **3.15** according to the general method A. The crude product was dissolved in MeOH containing 10% Pd/C (4 mg), and the mixture was purged four times with  $H_2$ . The suspension was stirred under  $H_2$  (1 atm) for 2 h. The mixture was filtered through a

pad of celite, and the pad was washed with MeOH (5.0 mL). The combined filtrate and washings were concentrated under reduced pressure to yield 14 mg (72%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.31-7.24 (comp, 5 H), 7.19-7.14 (comp, 4 H), 4.60 (dd,  $J = 13.3, 4.8$  Hz, 2 H), 4.47 (dd,  $J = 15.2, 5.1$  Hz, 1 H), 4.35 (dd,  $J = 13.3$  Hz, 4.8 Hz, 1 H), 2.97 (dd,  $J = 14.4, 5.1$  Hz, 1 H), 2.86 (dd,  $J = 14.4, 5.1$  Hz, 1 H), 2.78-2.71 (comp, 2 H), 2.00-1.92 (comp, 2 H), 1.82 (s, 3 H), 1.76-1.69 (comp, 2 H), 1.56-1.46 (br, 3 H), 1.27-1.18 (br, 3 H);  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.7, 175.6, 174.4, 174.3, 173.6, 139.8, 134.1, 131.3, 129.4, 128.4, 128.1, 121.3, 61.3, 61.2, 56.3, 52.3, 44.1, 37.1, 36.7, 33.0, 32.60, 26.2, 22.2, mass spectrum (ESI)  $m/z$  654.23004 [ $\text{C}_{29}\text{H}_{44}\text{N}_5\text{O}_9\text{P Na}$  (M+1) requires 654.2300].

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.31-7.24 (comp, 5 H, C7-H, or C22-H, C23-H or C24-H), 7.19-7.14 (comp, 4 H, C8-H or C22-H or C23-H or C24-H), 4.60 (dd,  $J = 13.3$  Hz, 4.8 Hz, 2 H, C20-H), 4.47 (dd,  $J = 15.2, 5.1$  Hz, 4.1 Hz, 1 H, C3-H), 4.35 (dd,  $J = 13.3, 4.8$  Hz, 1 H, C15-H), 2.97 (dd,  $J = 14.4, 5.1$  Hz, 1 H, C5-H), 2.86 (dd,  $J = 14.4, 5.1$  Hz, 1 H, C5-H), 2.78-2.71 (comp, 2 H, C16-H), 2.00-1.92 (comp, 2 H, C11-H), 1.82 (s, 3 H, C1-H), 1.76-1.69 (comp, 2 H, C11-H), 1.56-1.46 (br, 3 H, C13-H and C-12H), 1.27-1.18 (br, 3 H);  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.7 (C9 or C16), 175.6 (C9 or C16), 174.4 (C2, or C17 or C18), 174.3 (C2, or C17 or C18), 173.6 (C2, or C17 or C18), 139.8 (C9), 134.1 (C7), 131.3 (C6), 129.4 (C8), 128.4 (C-Ar), 128.1 (C-Ar), 121.3 (C-Ar), 61.3 (C3 or C10 or C15), 61.2 (C3 or C10 or C15), 56.3 (C3 or C10 or C15), 52.3, (C3 or C10 or C15) 44.1 (C20), 37.1 (C5 or C17), 36.7 (C5 or C17), 33.0

(C11), 32.60 (C11'), 26.2 (C12 or C12' or C13), 22.3 (C12 or C12' or C13), 22.2 (C12 or C12' or C13).

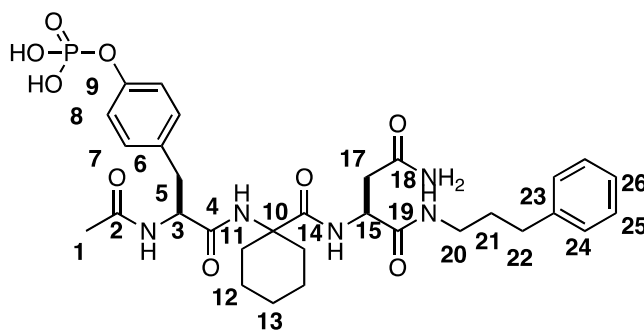


**4(*S*)-2-Acetamido-3-((1-(((*S*)-4-amino-1,4-dioxo-1-(phenethylamino)butan-2-yl)carbamoyl)cyclohexyl)amino)-3-oxopropyl)phenyl dihydrogen phosphate (3.07).**

Prepared from **3.24** according to the representative procedure for the synthesis of tripeptides to yield 27 mg (34%) of the title compound as over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.27 – 7.24 (comp, 6 H), 7.17-7.14 (comp, 3 H), 4.65 (dd,  $J$  = 9.5, 6.3 Hz, 1 H), 4.53 (dd,  $J$  = 6.9, 5.1 Hz, 1 H), 3.42-3.38 (comp, 2 H), 3.13-3.09 (comp, 1 H), 2.94-2.86 (comp, 1 H), 2.84–2.68 (comp, 4H), 1.98-1.95 (br, 2 H), 1.88 (s, 3 H), 1.8-1.71 (br, 3 H), 1.54 (comp, 3 H), 1.4-1.2 (m, 4 H);  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.6, 175.7, 174.2, 173.6, 172.9, 140.5, 134.7, 131.4, 129.8, 129.4, 127.3, 121.4, 61.3, 56.3, 52.1, 42.4, 37.2, 36.7, 36.5, 32.8, 26.2, 22.3. mass spectrum (ESI)  $m/z$  668.2456 [ $\text{C}_{30}\text{H}_{39}\text{N}_5\text{O}_9\text{P}$ -1 ( $\text{M}+1$ ) requires 668.2464].

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.27 – 7.24 (comp, 6 H, C7-H, or C21-H, C22-H or C23-H), 7.17-7.14 (comp, 3 H, C8-H and C25-H), 4.65 (dd,  $J$  = 9.5, 6.3 Hz, 1 H, C3-H), 4.53 (dd,  $J$  = 6.9, 5.1 Hz, 1 H, C14-H), 3.42-3.38 (comp, 2 H, C17-H), 3.13-3.09 (comp, 1 H, C5-H), 2.94-2.86 (comp, 1 H, C5-H), 2.84–2.68 (comp, 4

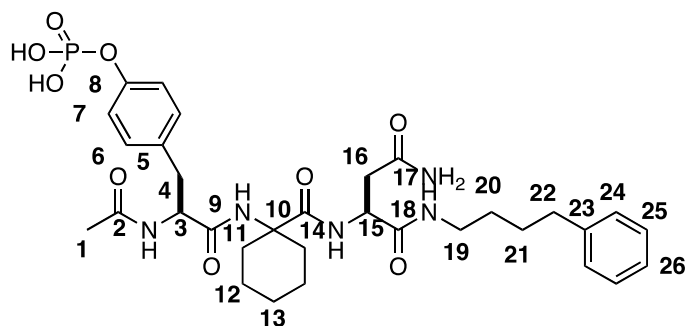
H, C21–H and C17–H) 1.98–1.95 (comp, 2 H, C11–H), 1.88 (s, 3 H, C1–H), 1.8–1.71 (br, 2 H, C11'–H), 1.54–1.44 (br, 3 H, C12–H and C13–H), 1.4–1.2 (br, 4 H C12'–H and C13–H).  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.6 (C9 or C16), 175.7 (C9 or C16), 174.2 (C2, C17, or C18), 173.6 (C2, C17, or C18), 172.9 (C2, C17, or C18), 140.5 (C–Ar), 134.7 (C–Ar), 131.4 (C–Ar), 129.8 (C–Ar), 129.4 (C–Ar), 127.3 (C–Ar), 121.4 (C–Ar), 61.3 (C3, C10, or C15), 56.3 (C3, C10, or C15), 52.1 (C3, C10, or C15), 42.4 (C20), 37.2 (C5 or C17), 36.7 (C5 or C17), 32.8 (2) (C11 and C11'), 26.2 (C13), 22.32 (C12), 22.23 (C12').



**4(S)-2-Acetamido-3-((1-(((S)-4-amino-1,4-dioxo-1-((3-phenylpropyl)amino)butan-2-yl)carbamoyl)cyclohexyl)amino)-3-oxopropyl)phenyl dihydrogen phosphate (3.08).** Prepared from **3.25** according to the representative procedure for the synthesis of tripeptides to yield 9 mg (28%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.24 – 7.19 (comp, 6 H), 7.14–7.12 (comp, 3 H), 4.65 (dd,  $J$  = 9.5, 6.3 Hz, 1 H), 4.53 (dd,  $J$  = 6.9, 5.1 Hz, 1 H), 3.26–3.22 (comp, 2 H), 3.13–3.11 (comp, 1 H), 2.92–2.71 (comp, 3 H) 2.66–2.62 (t,  $J$  = 7.5 Hz, 2 H), 2.05–1.93 (br s, 2 H), 1.88 (s, 3 H), 1.8–1.71 (comp, 3 H), 1.54 (br s, 1 H),

1.23 (br s, 3 H);  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  175.5, 174.5, 173.1, 172.4, 171.8, 151.6, 139.3, 133.1, 130.1, 128.6, 128.3, 126.1, 120.2, 60.0, 55.2, 51.0, 41.2, 36.0, 35.6, 35.3, 31.6 (2), 25.0, 21.1. mass spectrum (ESI)  $m/z$  660.2793 [ $\text{C}_{31}\text{H}_{43}\text{N}_5\text{O}_9\text{P}-1$  (M+1) requires 660.2793].

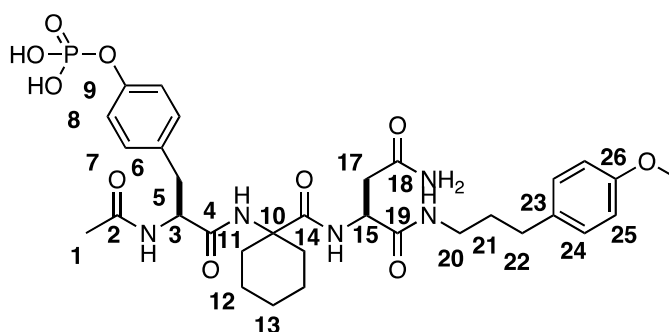
**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.24 – 7.19 (comp, 6 H, C7-H, or C24-H, C25-H or C26-H), 7.14-7.12 (comp, 3 H, C8-H and C26-H), 4.65 (dd,  $J$  = 9.5, 6.3 Hz, 1 H, C3-H), 4.53 (dd,  $J$  = 6.9, 5.1 Hz, 1 H, C15-H), 3.26-3.22 (comp, 2 H, C20-H), 3.13–3.11 (comp, 1 H, C4-H), 2.92–2.71 (comp, 3 H, C22-H and C4-H) 2.66-2.62 (t,  $J$  = 7.5 Hz, 2 H, C21-H), 2.05-1.93 (br s, 2 H, C11-H), 1.88 (s, 3 H, C1-H), 1.8-1.71 (m, 3.6 H, C11'-H), 1.54 (br s, 3 H, C12-H and C13-H), 1.23 (br s, 3 H, C12'-H and C13-H ).  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  175.5 (C4 or C14), 174.5 (C4 or C14), 173.1 (C2, C18, or C19), 172.4 (C2, C18, or C19), 171.8 (C2, C18, or C19), 151.6 (C-Ar), 139.3 (C-Ar), 133.1(C-Ar), 130.1 (C-Ar), 128.6 (C-Ar), 128.3 (C-Ar), 126.1 (C-Ar), 120.2(C-Ar), 60.0 (C3, C10, C15), 55.2 (C3, C10, C15), 51.0 (C3, C10, C15), 41.2 (C20), 36.0(C22), 35.6 (C5 or C17), 35.3 (C5 or C17), 31.6 (2) (C11 and C11'), 25.0 (C21), 21.1 (C12' and C13).



**4(S)-2-Acetamido-3-((1-(((S)-4-amino-1,4-dioxo-1-((4-phenylbutyl)amino)butan-2-yl)carbamoyl)cyclohexyl)amino)-3-oxopropyl)phenyl dihydrogen phosphate (3.09).** Prepared from **3.26** according to the representative procedure for the synthesis of tripeptides to yield 20 mg (15%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.23 – 7.19 (comp, 6 H), 7.14-7.12 (comp, 3 H), 4.66 (dd,  $J = 9.5, 6.1$  Hz, 1 H), 4.53 (dd,  $J = 6.9, 5.1$  Hz, 1 H), 3.26-3.08 (comp, 3 H), 2.92-2.76 (comp, 3 H), 2.66-2.62 (t,  $J = 7.5$  Hz, 2 H), 2.05-1.93 (br s, 2 H), 1.88 (s, 3 H), 1.8-1.71 (comp, 3 H), 1.54 (br s, 3 H), 1.23 (br s, 3 H);  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.6, 175.6, 174.2, 173.6, 172.9, 143.7, 131.4, 129.5, 129.3, 126.7, 121.4, 61.2, 52.3, 40.4, 37.1, 36.7, 36.5, 33.0, 32.6, 29.9, 29.8, 26.2, 22.3, 22.2 mass spectrum (ESI)  $m/z$  672.2804 [ $\text{C}_{32}\text{H}_{44}\text{N}_5\text{O}_9\text{P}-1$  (M+1) requires 672.2802].

**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.23 – 7.19 (comp, 6 H, C7-H, or C24-H, C25-H or C26-H), 7.14-7.12 (comp, 3 H, C8-H and C26-H), 4.66 (dd,  $J = 9.5, 6.1$  Hz, 1 H, C3-H), 4.53 (dd,  $J = 6.9, 5.1$  Hz, 1 H, C15-H), 3.26-3.08 (comp, 3 H, C19-H), 2.92-2.76 (comp, 3 H, C22-H and C4-H), 2.66-2.62 (comp, 2 H, C20-H or C21-H), 2.05-1.93 (m, 2 H, C11-H), 1.88 (s, 3 H, C1-H), 1.8-1.71 (m, 3.6 H, C11'-H),

1.54 (br s, 3 H, C12–H and C13–H), 1.23 (br s, 3 H, C12'–H, C13–H).  $^{13}\text{C}$  (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.6 (C9 or C14), 175.6 (C9 or C14), 174.2 (C2, C17, or C18), 173.6 (C2, C17, or C18), 172.9 (C2, C17, or C18), 143.7 (C–Ar), 131.4 (C–Ar), 129.5 (C–Ar), 129.3 (C–Ar), 126.7 (C–Ar), 121.4 (C–Ar), 61.2 (C3–H, C10–H, or C15–H), 56.9 (C3–H, C10–H, or C15–H), 52.3 (C3–H, C10–H, or C15–H), 40.4 (C19–H), 37.1 (C22), 36.7 (C4 or C16), 36.5 (C4 or C16), 33.0 (C11 or C11'), 32.6 (C11 or C11'), 29.9 (C20 or C21), 29.8 (C20 or C21), 26.2 (C13 and C12 or C12'), 22.3 (C13 and C12 or C12'), 22.2 (C1).

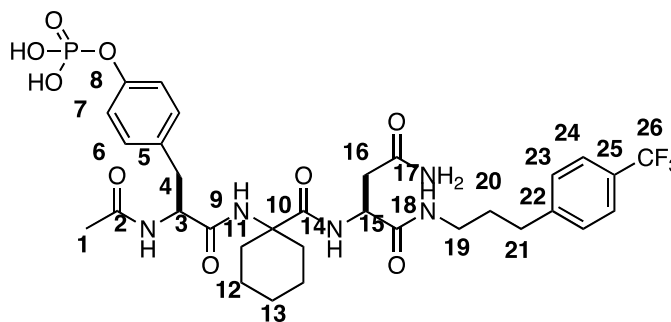


**4(*S*)-2-Acetamido-3-(1-((*S*)-4-amino-1-(3-(4-methoxyphenyl)propylamino)-1,4-dioxobutan-2-ylcarbamoyl)cyclohexylamino)-3-oxopropyl)phenyl dihydrogen phosphate (3.48)** Prepared from **3.59** according to the representative procedure for the synthesis of tripeptides to yield 15 mg (20%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.22 (d,  $J$  = 8.2 Hz, 2 H), 7.14-7.10 (comp, 4 H), 6.77 (d,  $J$  = 8.2 Hz, 2 H), 4.65 (dd, 1.0 H,  $J$  = 9.2, 6.2 Hz, 1 H), 4.54-4.50 (comp, 1 H), 3.73 (s, 3 H), 3.33-3.27 (comp, 2 H), 3.13-3.10 (dd,  $J$  = 14.0, 6.2 Hz, 1 H), 2.91-2.75 (comp, 3 H), 2.57 (t,  $J$  = 7.52 Hz, 2 H), 2.00-1.88 (br s, 2 H), 1.88 (s, 3 H), 1.82-1.71 (comp, 5 H), 1.56-1.54 (comp, 4 H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.6,



175.7, 174.3, 173.6, 172.9, 159.3, 151.7, 135.2, 134.7, 131.4, 130.4, 121.3, 114.8, 61.2, 56.3, 55.6, 52.3, 40.3, 37.1, 36.6, 33.1, 32.4, 26.2, 22.3, 22.2. mass spectrum (ESI)  $m/z$  638.2963 [ $C_{32}H_{45}N_5O_{10}P-1$  (M+1) requires 638.2960].

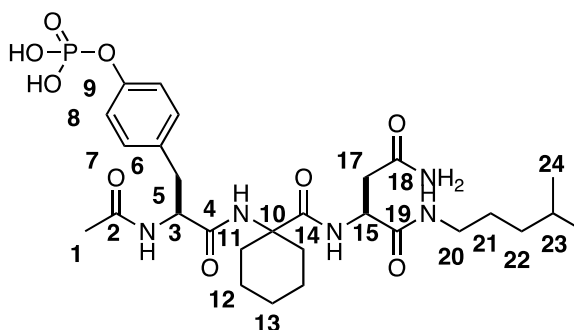
**NMR Assignments.**  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta$  7.22 (d,  $J = 8.2$  Hz, 2 H, C24-H), 7.14-7.10 (comp, 2 H, C7-H and C8-H), 6.77 (d,  $J = 8.2$  Hz, 2 H, C25-H), 4.65 (dd, 1 H,  $J = 9.2, 6.2$  Hz, C3-H), 4.54-4.50 (comp, 1 H, C15-H), 3.73 (s, 3 H, C27-H), 3.23 (dd,  $J = 13.1, 6.8$  Hz, 2 H, C20-H), 3.12 (dd,  $J = 14.0, 6.2$  Hz, 1 H, C5-H), 2.91-2.75 (comp, 3 H C5-H and C22-H), 2.57 (t,  $J = 7.52$  Hz, 2 H, C21-H), 2.00-1.88 (comp, 3 H, C11-H), 1.88 (s, 3 H, C1-H), 1.82-1.71 (comp, 5 H, C11'-H), 1.56-1.54 (comp, 3 H, C13-H and C12-H or C12'-H), 1.23 (br s, 3 H, C13-H and C12-H or C12'-H).  $^{13}C$  NMR (125 MHz,  $CD_3OD$ )  $\delta$  176.6 (C9 or C14), 175.7 (C9 or C14), 174.3 (C2, C17, or C18), 173.6 (C2, C17, or C18), 172.9 (C2, C17, or C18), 159.3 (C-Ar), 151.7 (C-Ar), 135.2 (C-Ar), 134.7 (C-Ar), 131.4 (C-Ar), 130.4 (C-Ar), 121.3 (C-Ar), 114.8 (C-Ar), 61.2 (C3-H, C10-H, or C15-H), 56.3 (C3-H, C10-H, or C15-H), 55.6 (C27), 52.3 (C3-H, C10-H, or C15-H), 40.3 (C20), 37.1 (C22), 36.6 (C5 or C17), 33.1 (C3-H, C10-H, or C15-H), 32.4 (C11), 26.2 (C21), 22.3 (C12) 22.2 (C13 and C1).



**4(*S*)-2-Acetamido-3-(1-((*S*)-4-amino-1-(3-(4-methoxyphenyl)propylamino)-1,4-dioxobutan-2-ylcarbamoyl)cyclohexylamino)-3-oxopropyl)phenyl dihydrogen phosphate (3.47).** Prepared from **3.58** according to the representative procedure for the synthesis of tripeptides to yield 31 mg (94%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.47 (d,  $J$  = 8.2 Hz, 2 H), 7.36 (d,  $J$  = 8.0 Hz, 2 H), 7.22 (d,  $J$  = 8.2 Hz, 2 H), 7.14 (d,  $J$  = 8.0 Hz, 2 H), 4.65 (dd,  $J$  = 8.5, 7.5 Hz, 1 H), 4.53-4.51 (m, 1 H), 3.22 (t,  $J$  = 6.59 Hz, 2 H), 3.11 (dd,  $J$  = 14.0, 6.1 Hz, 1 H), 2.91-2.75 (comp, 3 H), 2.00-1.88 (comp, 3 H), 1.88 (s, 3 H), 1.82-1.71 (comp, 5 H), 1.56-1.54 (comp, 4 H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.6, 175.7, 174.3, 173.6, 172.9, 159.3, 151.7, 135.2, 134.7, 131.4, 130.4, 121.3, 114.8, 61.2, 56.3, 55.6, 52.3, 40.3, 37.1, 36.6, 33.1, 32.4, 26.2, 22.3, 22.2. mass spectrum (ESI)  $m/z$  726.2527 [ $\text{C}_{32}\text{H}_{40}\text{N}_5\text{O}_9\text{P}$ -1 ( $\text{M}$ -1) requires 726.2521].

**NMR Assignment.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.47 (d,  $J$  = 8.2 Hz, 2 H, C24-H), 7.36 (d,  $J$  = 8.0 Hz, 2 H, C6-H), 7.22 (d,  $J$  = 8.2 Hz, 2 H, C23-H), 7.14 (d,  $J$  = 8.0 Hz, 2 H, C7-H), 4.65 (dd,  $J$  = 8.5, 7.5 Hz, 1 H, C3-H), 4.53-4.51 (m, 1 H, C15-H), 3.22 (t,  $J$  = 6.59 Hz, 2 H, C19-H), 3.13-3.10 (dd,  $J$  = 14.0, 6.1 Hz, 1 H, C4-H), 2.91-2.75 (comp, 3 H C21-H and C4-H), 2.00-1.88 (comp, 2 H, C11-H), 1.88 (s, 3 H, C1-H), 1.82-

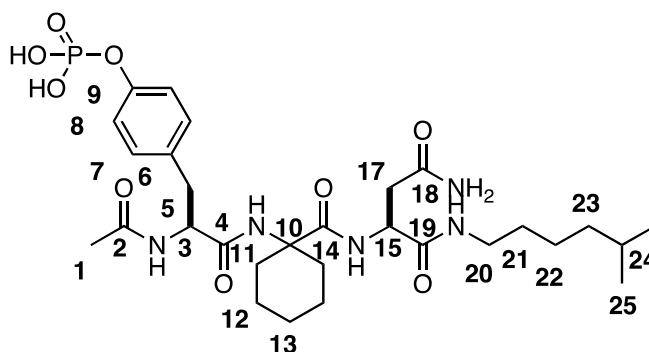
1.71 (comp, 2 H C11'-H), 1.56-1.54 (comp, 4 H, C13 and C12 or C12'), 1.23 (comp, 3 H, C13-H and C12-H or C12'-H); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD) δ 176.6 (C9 or C14), 175.7 (C9 or C14), 174.3 (C2, C17, or C18), 173.6 (C2, C17, or C18), 172.9 (C2, C17, or C18), 159.3 (C8), 151.7 (C-Ar), 135.2 (C-Ar), 134.7 (C-Ar), 131.4 (C-Ar), 130.4 (C-Ar), 121.3 (C-Ar), 114.8 (C-Ar), 61.2 (C3-H, C10-H, or C15-H), 56.3 (C3-H, C10-H, or C15-H), 55.6 (C3-H, C10-H, or C15-H), 52.3 (C19), 40.3 (C21), 37.1 (C22), 36.6 (C4 or C16), 33.1 (C4 or C16), 32.4 (2) (C11 and C11'), 26.2 (C20), 22.3 (C13), 22.2 (C1 and C12).



**4-((*S*)-2-acetamido-3-((1-(((*S*)-4-amino-1-((4-methylpentyl)amino)-1,4-dioxobutan-2-yl)carbamoyl)cyclohexyl)amino)-3-oxopropyl)phenyl dihydrogen phosphate (3.43).** Prepared from **3.55** and **3.15** according to the representative procedure for the synthesis of tripeptides to yield 12 mg (87%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.26 (d,  $J = 8.2$  Hz, 2 H), 7.14 (d,  $J = 8.2$  Hz, 2 H), 4.62-4.68 (m, 1 H), 4.47-54 (m, 1 H), 3.08-3.20 (comp, 3 H), 2.91 (dd,  $J = 9.2$  Hz, 7.2 Hz, 1 H), 2.80 (dd,  $J = 12.3$  Hz, 7.0 Hz, 1 H), 2.73 (dd,  $J = 12.3$  Hz, 7.0 Hz, 1 H), 2.00-1.92 (comp, 2 H), 1.91 (s, 3 H), 1.74-1.61 (comp, 2 H), 1.52-1.49 (comp, 5 H), 1.38 (m, 1 H) 1.22-1.17 (comp, 4 H), 0.87 (d,  $J = 6.5$  Hz, 6 H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ ) 176.6, 175.7, 174.3, 173.6, 152.0, 134.6, 131.4, 121.4, 61.2, 56.4, 52.3, 41.0, 37.14, 37.05, 33.0, 32.5, 29.0, 28.2, 26.2, 23.0, 22.3, 22.2. mass spectrum (ESI)  $m/z$  624.2806 [ $\text{C}_{28}\text{H}_{43}\text{N}_5\text{O}_{10}\text{NaP}+1$  ( $\text{M}+1$ ) requires 624.2804].

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.26 (d,  $J = 8.2$  Hz, 2 H, C7-H), 7.14 (d,  $J = 8.2$  Hz, 2 H, C8-H), 4.65 (comp, 1 H, C3-H), 4.51 (comp, 1 H, C15-H), 3.16 (m, 3 H, C20-H and C5-H), 2.91 (dd,  $J = 9.2$  Hz, 7.2 Hz, 1 H, C5-H), 2.80 (dd,  $J = 12.3$  Hz, 7.0 Hz, 1 H, C17-H), 2.73 (dd,  $J = 12.3$  Hz, 7.0 Hz, 1 H, C17-H), 2.00-1.92

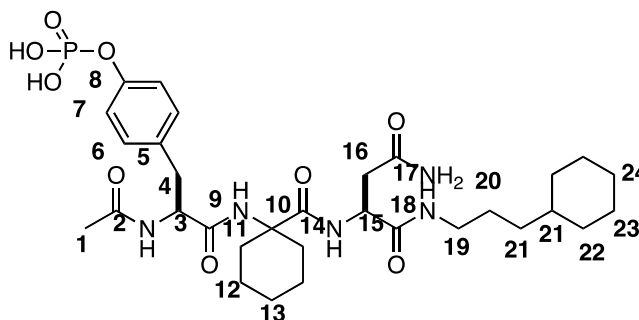
(comp, 2 H, C11–H), 1.91 (s, C1–H), 1.74–1.61 (comp, 2 H, C11'–H), 1.52–1.49 (br s, 3 H, C13–H and C12–H or C12'–H ), 1.21–1.19 (br s, 3 H, C13–H and C12–H or C12'–H );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  176.58 (C4 or C14), 175.5 (C4 or C14), 174.3 (C2, C18, or C19), 173.6 (C2, C18, or C19), 172.8 (C2, C18, or C19), 152.0 (C9), 134.6 (C7), 131.4 (C6), 121.4 (8), 61.2 (C3, C10 or C15), 56.4 (C3, C10 or C15), 52.3 (C3, C10 or C15), 41.0 (C20), 37.14 (C5 or C17), 37.10 (C5 or C17), 36.8 (C11 or C11'), 33.0 (C11 or C11'), 32.5 (C21, C22, C23, C24, or C25), 29.0 (C21, C22, C23, C24, or C25), 28.2 (C21, C22, C23, C24, or C25), 26.1 (C21, C22, C23, C24, or C25), 23.0 (C21, C22, C23, C24, or C25), 22.3 (C12), 22.2 (C12'), 22.19 (C1).



**4-((*S*)-2-Acetamido-3-((1-(((*S*)-4-amino-1-((5-methylhexyl)amino)-1,4-dioxobutan-2-yl)carbamoyl)cyclohexyl)amino)-3-oxopropyl)phenyl dihydrogen phosphate (3.44).** Prepared from **3.54** according to the representative procedure for the synthesis of tripeptides to yield 13 mg (22%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95% B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.19 (d,  $J = 8.2$  Hz, 2 H), 7.05 (d,  $J = 8.2$  Hz, 2 H), 4.57–4.55 (dd, 1 H,  $J = 8.9, 6.2$  Hz, 1 H), 4.54–4.48 (dd,  $J = 7.2, 5.5$  Hz, 1 H), 3.10–3.01 (comp, 3 H), 2.85–2.79 (comp, 1 H), 2.74–2.61 (comp, 2 H), 2.00–1.88 (comp, 5

H), 1.74-1.61 (comp, 2 H), 1.52-1.49 (comp, 6 H), 1.29-1.12 (comp, 7 H) 0.78-0.76 (comp, 6 H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ ) 176.6, 175.7, 174.3, 173.6, 152.0, 134.643, 131.4, 121.4, 61.2, 56.4, 52.3, 41.1, 37.1, 37.0, 33.0, 32.5, 29.0, 28.2, 26.2, 23.0, 22.2, 22.2 mass spectrum (ESI)  $m/z$  638.2963 [ $\text{C}_{29}\text{H}_{45}\text{N}_5\text{O}_9\text{P}-1$  (M+1) requires 638.2960.]

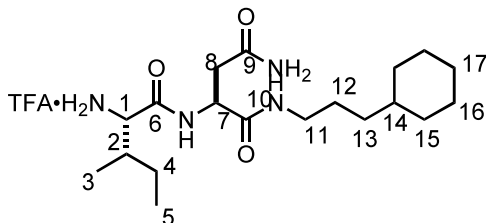
**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.19 (d,  $J = 8.2$  Hz, 2 H, C6-H), 7.05 (d,  $J = 8.2$  Hz, 2 H, C7-H), 4.57-4.55 (dd, 1 H, 8.9 Hz, 6.2 Hz, C3-H), 4.54-4.48 (m, 1 H, C15-H), 3.10-3.01 (comp, 3 H, C19-H and C4-H), 2.85-2.79 (m, 1 H, C4-H), 2.74-2.61 (comp, 2 H, C16-H), 2.00-1.88 (comp, 5 H, C11-H and C1-H), 1.74-1.61 (comp, 2 H, C11-H), 1.52-1.49 (comp, 6.0 H, C12-H, C13-H, C20-H, C21-H, C22-H or C23-H), 1.29-1.12 (comp, 7 H, C12-H, C13-H, C20-H, C21-H, C22-H or C23-H) 0.78-0.76 (m, 6.0 H, C24-H).



**4-((S)-2-Acetamido-3-((1-(((S)-4-amino-1-((3-cyclohexylpropyl)amino)-1,4-dioxobutan-2-yl)carbamoyl)cyclohexyl)amino)-3-oxopropyl)phenyl dihydrogen phosphate (3.63).** Prepared from **3.53** according to the representative procedure for the synthesis of tripeptides to yield 19 mg (75%) of the title compound over two steps. The crude material was purified via preparative RP HPLC using a gradient of 10% B to 95%

B over 30 min.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.28-7.26 (comp, 2 H), 7.15-7.13 (comp, 2 H), 4.68-4.64 (comp, 1 H), 4.54-4.48 (comp, 1 H), 3.19-3.07 (comp, 3 H), 2.92-2.89 (dd,  $J = 13.7$  Hz, 9.6 Hz, 1 H), 2.83-2.79 (comp, 2 H), 2.00-1.92 (br d, 2 H), 1.88 (s, 3 H), 1.74-1.61 (comp, 7.0 H), 1.59-1.49 (comp, 5 H), 0.87-0.85 (comp, 2 H), 1.24-1.08 (comp, 9 H) 0.95-0.81 (comp, 2 H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ ) 177.6, 176.58, 175.6, 174.3, 173.7, 173.0, 172.8, 152.18, 134.28, 131.4, 131.3, 121.4, 61.3, 61.2, 60.2, 56.4, 55.7, 52.2, 41.0, 38.8, 38.0, 37.0, 36.7, 35.7, 34.5, 33.3, 31.8, 32.4, 27.77, 27.68, 26.4, 26.2, 22.45, 22.42, 22.37, 22.31, 22.2. mass spectrum (ESI)  $m/z$  664.3120 [ $\text{C}_{31}\text{H}_{47}\text{N}_5\text{O}_9\text{P}$ -1 (M+1) requires 664.3117]

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.28-7.24 (comp, 2 H, C6-H), 7.15-7.13 (comp, 2 H, C7-H), 4.68-4.64 (m, 1 H, C3-H), 4.54-4.48 (m, 1 H, C15-H), 3.19-3.07 (comp, 2 H, C19-H and C4-H), 2.92-2.89 (dd,  $J = 13.7$ , 9.6 Hz, 1 H, C4-H), 2.83-2.79 (comp, 2 H, C16-H), 2.00-1.92 (comp, 2 H, C11-H), 1.88 (s, 3 H, C1-H), 1.74-1.61 (comp, 4 H, C11-H and C23-H), 1.52-1.49 (comp, 5 H, C20-H, C21-H and C22-H), 0.87-0.85 (comp, 2 H, C24-H or C25-H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ ) 177.6 (C9 or C14), 175.6 (C9 or C14), 174.3 (C2, C17, or C18), 173.7 (C2, C17, or C18), 173.0 (C2, C17, or C18), 172.8 (C2, C17, or C18), 152.18 (C8), 134.28 (C6), 131.4 (C5), 121.4 (C7), 61.3 (C3, C10, or C15), 61.2 (C3, C10, or C15), 60.2 (C3, C10, or C15), 56.4, 55.7, 52.2, 41.0 (C19), 38.8 (C4 or C16), 38.0 (C4 or C16), 37.0 (C11), 36.7 (C11'), 35.7 (C22, C23, C24, C21, or C20) 34.5 (C22, C23, C24, C21, or C20), 33.3 (C22, C23, C24, C21, or C20), 31.8 (C22, C23, C24, C21, or C20), 32.4 (C22, C23, C24, C21, or C20), 27.77 (C22, C23, C24, C21, or C20), 26.4, (C22, C23, C24, C21, or C20) 22.45, (C22, C23, C24, C21, or C20), 22.31 (C22, C23, C24, C21, or C20) 22.2 (C1).

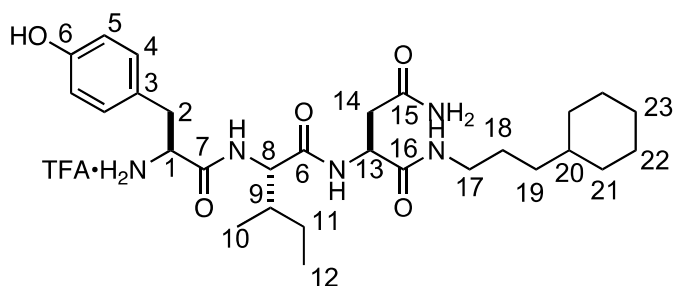


***N*-TFA-Ile-Asn-*N*-(propylCyHex) (3.71)** Prepared from **3.70** and Boc-Ile according to the general method B to give the crude product as a white solid: mp = 182-183 °C. The white solid was then dissolved in CH<sub>2</sub>Cl<sub>2</sub> (2.0 mL) whereupon TFA (530 mg, 4.6 mmol) was added to the mixture and the resulting solution was allowed to stir for 8 h at room temperature. It was then concentrated to dryness under reduced pressure to give a brown solid. The crude solid was triturated with Et<sub>2</sub>O to provide 252 mg (46%, 2 steps) of the title compound as a white powder. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 4.73 (comp, 1 H), 3.70 (comp, 1 H), 3.18-3.08 (comp, 2 H), 2.74-2.61 (comp, 2 H), 1.94-1.90 (m, 1 H), 1.72-1.63 (comp, 5 H), 1.60-1.55 (comp, 1H), 1.49 (p, *J* = 8.0 Hz, 2 H), 1.29-1.14 (comp, 7 H), 1.01 (d, *J* = 7.3 Hz, 3 H), 0.96 (t, *J* = 7.3 Hz, 3 H), 0.92-0.84 (comp, 2 H); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD) δ 174.5, 172.3, 169.3, 58.9, 51.7, 40.9, 38.7, 38.0, 37.8, 35.7, 34.5, 27.7, 27.5, 25.4, 11.6 mass spectrum (ESI) *m/z* 369.2861 [C<sub>19</sub>H<sub>37</sub>N<sub>4</sub>O<sub>3</sub>(M+1) requires 369.2860].

**NMR Assignments:** <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 4.73 (comp, 1 H, C7-H), 3.70 (comp, 1 H, C1-H), 3.18-3.08 (comp, 2 H, C11-H), 2.74-2.61 (comp, 2 H, C8-H), 1.94-1.90 (m, 1 H, C2-H), 1.72-1.63 (comp, 5 H, C12-H, C13-H, C14-H, C15-H, C16-H or C17-H), 1.60-1.55 (m, 1H, C12, C13, C14, C15, C16 or C17-H), 1.49 (p, *J* = 8.0 Hz, 2 H, C4-H3), 1.29-1.14 (comp, 7 H, C12-H, C13-H, C14-H, C15-H, C16-H or C17-H), 1.01 (d, *J* = 7.3 Hz, 3 H, C3-H), 0.96 (t, *J* = 7.3 Hz, 3 H, C5-H), 0.92-0.84 (comp, 2 H,

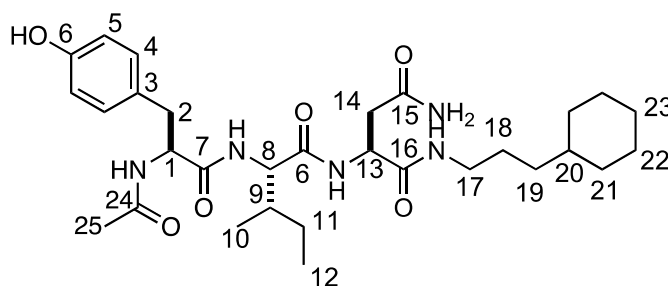


C12, C13, C14, C15, C16 or C17-H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  174.5 (C9), 172.3 (C10), 169.3 (C6), 58.9 (C1), 51.7 (C7), 40.9 (C11), 38.7 (C8), 38.0 (C2), 37.8 (C13, C15, or C15'), 35.7 (C13, C15, or C15'), 34.5 (C13, C15, or C15'), 27.7 (C18, C22, or C23), 27.5 (C4), 25.4 (C3), 11.6 (C5).



***N*-TFA-Tyr-Ile-Asn-*N*-(propyl-CyHex) (3.72).** Prepared from **3.71** and Boc-Tyr according to the general method B to give the crude product as a white solid: mp = 206–208 °C. The white solid was then dissolved in  $\text{CH}_2\text{Cl}_2$  (2.0 mL) whereupon TFA (530 mg, 4.6 mmol) was added to the mixture and the resulting solution was allowed to stir for 8 h at room temperature. It was then concentrated to dryness under reduced pressure to give a brown solid. The crude solid was triturated with  $\text{Et}_2\text{O}$  to provide 178 mg (66%, 2 steps) of the title product as a clear glass.  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.11-7.09 (comp, 2 H), 6.79-6.76 (comp, 2H), 4.67 (t,  $J$  = 6.6 Hz, 1 H), 4.22 (d,  $J$  = 7.5 Hz, 1 H), 4.1 (m, 1 H), 3.20-3.13 (comp, 3 H) 2.93-2.87 (m, 1 H) 2.73-2.64 (comp, 2 H), 1.86-1.82 (comp, 1 H), 1.72-1.62 (comp, 5 H), 1.58-1.55 (m, 1H), 1.49 (p,  $J$  = 7.6 Hz, 2 H), 1.29-1.14 (comp, 7 H), 0.94 (d,  $J$  = 7.5 Hz, 3 H), 0.91 (t,  $J$  = 7.5 Hz, 3 H), 0.92-0.84 (comp, 2 H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  174.7, 172.6, 172.5, 170.0, 158.3, 131.6, 125.9, 116.9, 116.3, 59.6, 55.7, 51.7, 40.9, 38.7, 38.2, 37.87, 37.80, 35.7, 34.5, 27.8, 27.5, 26.0, 15.9, 11.5. mass spectrum (ESI)  $m/z$  532.3497 [ $\text{C}_{28}\text{H}_{46}\text{N}_5\text{O}_3(\text{M}+1)$  requires 532.3499].

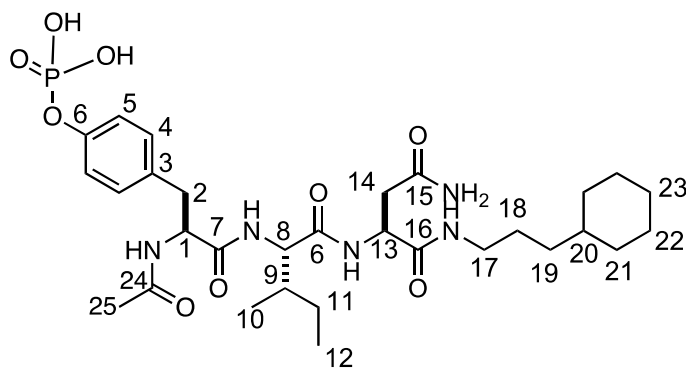
**NMR Assignments:**  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.11-7.09 (comp, 2 H, C4-H), 6.79-6.76 (comp, 2H, C5-H), 4.67 (t,  $J = 6.6$  Hz, 1.0 H, C1-H), 4.22 (d,  $J = 7.5$  Hz, 1 H, C13-H), 4.1 (m, 1 H, C8-H), 3.20-3.13 (comp, 3 H, C17-H and C2-H) 2.93-2.87 (m, 1 H, C2-H) 2.73-2.64 (comp, 2 H, C14-H), 1.86-1.82 (m, 1 H, C9-H), 1.72-1.62 (comp, 5 H, C18-H or C19-H or C20-H or C21-H or C22-H or C23-H), 1.58-1.55 (comp, 1H, C18 or C19-H or C20-H or C21-H or C22-H or C23-H), 1.49 (p,  $J = 7.6$  Hz, 2 H, C11-H), 1.29-1.14 (comp, 7 H, C18 or C19 or C20 or C21 or C22 or C23-H), 0.94 (d,  $J = 7.5$  Hz, 3 H, C10-H), 0.91 (t,  $J = 7.5$  Hz, 3 HC12-H), 0.92-0.84 (comp, 2 H, C18 or C19 or C20 or C21 or C22 or C23-H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  174.7 (C15), 172.6 (C16), 172.5 (C1), 170.0 (C6), 158.3 (C6), 131.6 (C4), 125.9 (C3), 116.9 (C5), 116.3 (C5), 59.6 (C1), 55.7 (C13), 51.7 (C8), 40.9 (C14), 38.7 (C2), 38.2 (C17), 37.9 (C19 or C21 or C21'), 37.8, (C19 or C21 or C21') 35.7 (C19 or C21 or C21'), 34.5 (C18, C20 or C22), 27.8 (C18, C20 or C22), 27.5 (C18, C20 or C22), 26.0 (C11), 15.9 (C10), 11.5 (C12).



***N*-Acyl-Tyr-Ile-Asn-*N*-(propyl-CyHex) (3.73)** Prepared from **3.72** and AcOH according to the general method B to yield 53 mg of the crude product (63%) as a white solid: decomp at 252 °C.  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.06-7.05 (comp, 2 H), 6.70-6.68 (comp, 2 H), 4.63 (t,  $J = 6.7$  Hz, 1 H), 4.58-4.57 (m, 1 H), 4.11 (d,  $J = 7.0$  Hz, 1 H), 3.19-3.10 (comp, 2 H) 3.06-3.02 (m, 1 H), 2.80-2.76 (m, 1H), 2.73-2.64 (comp, 2 H),

1.90 (s, 3 H), 1.86-1.82 (m, 1 H), 1.72-1.62 (comp, 6 H), 1.54-1.49 (comp, 4 H), 1.29-1.14 (comp, 9 H), 0.93-0.85 (comp, 10 H);  $^{13}\text{C}$  NMR (150 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  171.8, 171.5, 170.4, 170.2, 169.2, 155.6, 130.0, 128.1, 114.7, 57.1, 54.1, 49.7, 36.8, 36.7, 36.6, 36.2, 33.9, 32.7, 26.3, 26.1, 25.8, 24.2, 22.4, 15.2, 11.1. mass spectrum (ESI)  $m/z$  573.3448 [ $\text{C}_{30}\text{H}_{46}\text{N}_5\text{O}_6$  (M-1) requires 573.3454].

**NMR assignment:**  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.06-7.05 (comp, 2 H, C4-H), 6.70-6.68 (comp, 2H, C5-H), 4.63 (t,  $J = 6.7$  Hz, 1.0 H, C1-H), 4.57 (comp, 1 H, C13-H), 4.11 (d,  $J = 7.0$  Hz, 1 H, C8-H), 3.19-3.10 (comp, 2H, C17-H) 3.06-3.02 (m, 1 H, C2-H), 2.80-2.76 (comp, 1H, C2-H), 2.73-2.64 (comp, 2 H, C14-H), 1.90 (s, 3 H, C25-H), 1.86-1.82 (comp, 1 H, C9-H), 1.72-1.62 (comp, 6 H, C18-H, C19-H or C20-H or C21-H or C22-H or C23-H), 1.54-1.49 (comp, 4 H), 1.29-1.14 (comp, 9 H, C11-H or C18-H or C19-H or C20-H or C21-H or C22-H or C23-H), 0.93-0.85 (comp, 10 H, C10-H or C12-H or C18-H or C19-H or C20-H or C21-H or C22 -H or C23-H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  171.8 (C15), 171.5 (C16), 170.4 (C6), 170.2 (C1), 169.2 (C24), 155.6 (C6), 130.0 (C4), 128.1 (C3), 114.7 (C5), 57.1 (C1), 54.1 (C13), 49.7 (C13), 36.8 (C19 or C21 or C21'), 36.7 (C19 or C21 or C21'), 36.6 (C19 or C21 or C21'), 36.2, 33.9 (C18, C20 or C22), 32.7 (C18, C20 or C22), 26.3 (C18, C20 or C22), 26.1, 25.8, 24.2 (C25), 22.4 (C11), 15.2 (C10), 11.1 (C12).



***N*-Acyl-*p*Tyr-Ile-Asn-*N*-(propyl-Cy) (3.75).** 1-H- Tetrazole (24 mg, 0.35 mmol) (785 mg, 11.2 mmol) and diisopropylphosphoramidite (94  $\mu$ L, 0.28 mmol) were added to a solution of **3.77** (40 mg, 0.070 mmol) in DMF (2 mL) at 0 °C, and the solution was stirred at 0 °C for 1 h and then at room temperature for 15 h. The solution was cooled to 0 °C, and 6 M *tert*-butyl hydroperoxide in decane (34 mg, 0.17 mL) was added. The resulting solution was stirred at 0 °C for 30 min and then at room temperature for 5 h, whereupon it was cooled to 0 °C and 5% aqueous NaHSO<sub>3</sub> (5 mL) was added. The solution was stirred at 0 °C for 30 min and then at room temperature for 2 h. The mixture was transferred to a separatory funnel containing H<sub>2</sub>O (20 mL), and the layers were separated. The aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 15 mL) whereupon the combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure. The residue was triturated with Et<sub>2</sub>O (4 x 5 mL) to yield 250 mg (20%) of the product as a white solid.

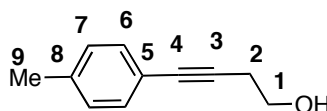
The product was dissolved in MeOH (0.5 mL) containing 10% Pd/C (1 mg, 10 wt %). The resulting mixture was purged four times with H<sub>2</sub>, and the suspension stirred under H<sub>2</sub> (1 atm) for 14 h. The mixture was filtered through a pad of celite and the pad was washed with MeOH. The combined filtrate and washings were concentrated to dryness under reduced pressure to yield 4 mg (99%) of the title compound as a clear

glass.  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.22 (d,  $J = 8.5$  Hz, 2 H), 7.13 (d,  $J = 8.5$  Hz, 2H), 4.65-4.60 (comp, 2 H), 4.13 (d,  $J = 7.0$  Hz, 1 H), 3.17-3.10 (comp, 3 H), 2.87-2.83 (m, 1 H) 2.70-2.68 (comp, 2 H), 1.86-1.82 (m, 1 H), 1.72-1.62 (comp, 5 H), 1.55-1.48 (comp, 3 H), 1.29-1.14 (comp, 8 H), 0.93-0.85 (comp, 9 H). mass spectrum (ESI)  $m/z$  652.3117 [ $\text{C}_{30}\text{H}_{47}\text{N}_5\text{O}_9\text{P}$  (M-1) requires 652.3190].

**NMR assignment:**  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.22 (d,  $J = 8.5$  Hz, 2 H, C4-H), 7.13 (d,  $J = 8.3$  Hz, 2 H, C5-H), 4.65-4.60 (comp, 2 H, C1-H and C13-H), 4.13 (d,  $J = 7.0$  Hz, 1 H, C-8-H), 3.17-3.10 (comp, 3 H, C2-H and C17-H), 2.87-2.83 (m, 1H, C2-H) 2.70-2.68 (comp, 2 H, C14-H), 1.86-1.82 (comp, 1 H, C9-H), 1.72-1.62 (comp, 5 H, C18-H, C19-H, C20-H, C21-H, C22-H, C23-H), 1.55-1.48 (comp, 3 H, C11-H and C20-H), 1.29-1.14 (comp, 8 H, C18-H, C19-H, C20-H, C21-H, C22-H, C23-H), 0.93-0.85 (comp, 9 H, C12-H and C10-H and C18-H or C19-H or C20-H or C21-H or C22-H or C23-H).

### General procedure for Sonogashira Cross Coupling.

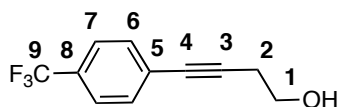
Triethylamine (2.9 g, 16.7 mmol) was added to a solution of the aryl iodide (3.3 mmol), alkynol (4.00 mmol), PdCl<sub>2</sub> (18 mg, 0.10 mmol), PPh<sub>3</sub> (80 mg, 0.3 mmol), and CuI (37.9 mg, 0.2 mmol) in MeCN (5.0 mL), and the mixture was stirred at 0 °C for 10 min. The ice bath was removed, and stirring was continued for 14 h at room temperature. The resulting mixture was diluted with Et<sub>2</sub>O (10 mL) then washed with brine (3 x 10 mL) and aqueous NH<sub>4</sub>Cl (2 x 10 mL). The organic layer was dried (MgSO<sub>4</sub>) and concentrated under reduced pressure, and purified by flash chromatography.



**4-(p-Tolyl)but-3-yn-1-ol (4.12).** Prepared from **4.04** and **4.24** according to the general procedure for Sonogashira cross coupling. The crude product was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 40 mg (74%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.30 (d, *J* = 8.1 Hz, 2 H), 7.09 (d, *J* = 8.1 Hz, 2 H), 3.80 (t, *J* = 6.2 Hz, 2 H), 2.68 (t, *J* = 6.2 Hz, 2 H), 2.34 (s, 3 H); <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 138.2, 131.8, 129.5, 120.4, 85.7, 82.8, 61.4, 24.1, 21.7. IR (neat) 3349, 2919, 2236, 1509 cm<sup>-1</sup>. mass spectrum (ESI) *m/z* 160.0888 [C<sub>11</sub>H<sub>12</sub>O requires 160.0888].

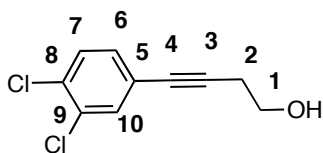
**NMR Assignment:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.30 (d, *J* = 7.9 Hz, 2 H, C6-H), 7.09 (d, *J* = 7.9 Hz, 2 H, C7-H), 3.80 (t, *J* = 6.2 Hz, 2 H, C1-H), 2.68 (t, *J* = 6.2 Hz, 2

H, C2-H), 2.34 (s, 3 H, C9-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  138.2 (C8), 131.8 (C7), 129.5 (C6), 120.4 (C5), 85.7 (C3), 82.8 (C4), 61.4 (C1), 24.1 (C2), 21.7 (C9).



**4-(4-(Trifluoromethyl)phenyl)but-3-yn-1-ol (4.11).** Prepared from **4.03** and **4.24** according to the general procedure for Sonogashira cross coupling. The crude product was purified by silica gel chromatography eluting with  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (99:1) to yield 87 mg (99%) of the title compound as a clear oil.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.53 (d,  $J$  = 8.9 Hz, 2 H), 7.49 (d,  $J$  = 8.9 Hz, 2 H), 3.84-3.80 (comp, 2 H), 2.71 (t,  $J$  = 6.5 Hz, 2 H);  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  132.1, 129.9 (q,  $J$  = 32.3 Hz), 127.5, 125.4, 124.5 (q,  $J$  = , 89.5, 81.4, 61.2, 24.0. IR (neat) 3349, 2890, 2237, 1615, 1406  $\text{cm}^{-1}$ . mass spectrum (CI)  $m/z$  214.0604 [ $\text{C}_{11}\text{H}_9\text{OF}_3$  requires 214.0605].

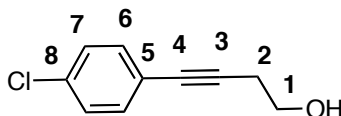
**NMR Assignment:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.53 (d,  $J$  = 8.9 Hz, 2 H, C7-H), 7.49 (d,  $J$  = 8.9 Hz, 2 H, C6-H), 3.84-3.80 (comp, 2 H, C1-H), 2.71 (t,  $J$  = 6.5 Hz, 2 H, C2-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  132.1 (C9), 129.9 (q,  $J$  = 32.3 Hz, C8), 127.5 (C6), 125.4 (C7), 122.8 (C5), 89.5 (C3), 81.4 (C4), 61.2 (C1), 24.0 (C2).



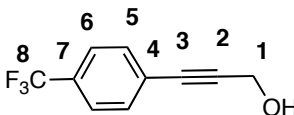
**4-(3,4-Dichlorophenyl)but-3-yn-1-ol (4.10).** Prepared from **4.03** and **4.24** according to the general procedure for Sonogashira cross coupling. The crude product

was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 56 mg (41%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.45 (d, *J* = 1.7 Hz, 1 H), 7.31 (d, *J* = 8.2 Hz, 1 H), 7.18 (dd, *J* = 8.2, 1.7 Hz, 1 H), 3.79 (t, *J* = 6.2 Hz, 2 H), 2.65 (t, *J* = 6.2 Hz, 2 H). <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 133.5, 132.6, 132.4, 131.0, 130.5, 123.6, 89.1, 80.3, 61.1, 23.9. IR (Neat) 3349, 2890, 2237, 1406, 1131 cm<sup>-1</sup>. mass spectrum (CI) *m/z* 213.9949 [C<sub>10</sub>H<sub>8</sub>OCl<sub>2</sub> requires 213.9952].

**NMR Assignment:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.45 (d, *J* = 1.7 Hz, 1 H, C10-H), 7.31 (d, *J* = 8.2 Hz, 1 H, C7-H), 7.18 (dd, *J* = 8.2, 1.7 Hz, 1 H, C6-H), 3.79 (t, *J* = 6.2 Hz, 2 H, C1-H), 2.65 (t, *J* = 6.2 Hz, 2 H, C2-H). <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 133.5 (C9), 132.6 (C6), 132.4 (C8), 131.0 (C7), 130.5 (C10), 123.6 (C5), 89.1 (C3), 80.3 (C4), 61.1 (C2), 23.9 (C1).



**4-(4-Chlorophenyl)but-3-yn-1-ol (4.09).** Prepared from **4.01** and **4.24** according to the general procedure for Sonogashira cross coupling. The crude product was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 87 mg (78%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.37 (d, *J* = 8.2 Hz, 2 H), 7.27 (d, *J* = 8.2 Hz, 2 H), 3.81 (t, *J* = 6.5 Hz, 2 H), 2.65 (t, *J* = 6.5 Hz, 2 H). The title compound was used in the subsequent hydrogenation without further purification.



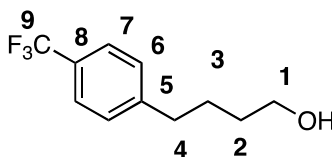


**3-(4-(Trifluoromethyl)phenyl)prop-2-yn-1-ol. (4.14)** Prepared from **4.06** and **4.23** according to the general procedure for Sonogashira cross coupling. The crude product was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 87 mg (92%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.37 (d, *J* = 8.2 Hz, 2 H), 7.27 (d, *J* = 8.2 Hz, 2 H), 4.52 (s, 2 H); <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 132.1, 130.6 (q, *J* = 32.1 Hz), 126.6, 125.5, 122.7, 89.9, 84.5, 51.5. IR (Neat) 3349, 2927, 2236, 1615, 1407 cm<sup>-1</sup>. mass spectrum (CI) *m/z* 200.0445 [C<sub>10</sub>H<sub>7</sub>OF<sub>3</sub> requires 200.0449].

**NMR Assignment:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.37 (d, *J* = 8.2 Hz, 2 H, C5-H), 7.27 (d, *J* = 8.2 Hz, 2 H, C6-H), 4.52 (s, 2 H, C1-H). <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 132.1 (C5), 130.6 (q, *J* = 32.1 Hz, C7), 126.6 (C6), 125.5 (C8), 122.7 (C4), 89.1 (C2), 84.3 (C3), 51.7 (C1).

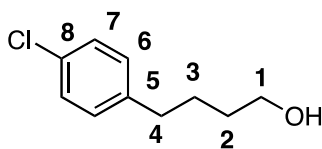
#### General procedure for the hydrogenation of the alkynols.

A solution of the alkynol (0.344 mmol) in MeOH (3.4 mL) containing 10% Pd/C (9.0 mg) or 5% PtO<sub>2</sub> (5.0 mg) was stirred under H<sub>2</sub> (1 atm) for 1 h. The catalyst was removed by filtration through a pad of celite, and the filtrate was concentrated under reduced pressure to yield the pure product.



**4-(4-(Trifluoromethyl)phenyl)butan-1-ol (4.18).** Prepared from **4.11** according to the general procedure for the hydrogenation of alkynols using Pd/C. The crude product was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 120 mg (59%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.52 (d, *J* = 8.2 Hz, 2 H), 7.27 (d, *J* = 8.2 Hz, 2 H), 3.67 (t, *J* = 6.2 Hz, 2 H), 2.71 (t, *J* = 7.5 Hz, 2 H), 1.76-1.69 (comp, 2 H), 1.64-1.57 (comp, 2 H). <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 146.7, 128.9, 128.3 (q, *J* = 32.3 Hz), 125.4, 124.6, 62.8, 35.7, 32.4, 27.5. IR (Neat) 3339, 2940, 2866, 1327, 1122 cm<sup>-1</sup>. mass spectrum (ESI) *m/z* 218.0915 [C<sub>11</sub>H<sub>13</sub>OF<sub>3</sub> requires 218.0918].

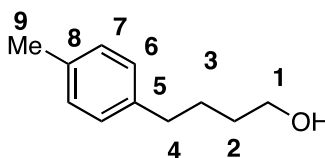
**NMR Assignments.** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.52 (d, *J* = 8.2 Hz, 2 H, C7-H), 7.27 (d, *J* = 8.2 Hz, 2 H, C6-H), 3.67 (t, *J* = 6.2 Hz, 2 H, C1-H), 2.71 (t, *J* = 7.5 Hz, 2 H, C4-H), 1.76-1.69 (comp, 2 H, C2-H or C3-H), 1.64-1.57 (comp, 2 H, C2-H or C3-H). <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 146.7 (C6), 128.3 (q, *J* = 32.3 Hz, C8), 128.3 (C9), 125.4 (C7), 124.6 (C5), 62.8 (C1), 35.7 (C4), 32.4 (C2 or C3), 27.5 (C2 or C3).



**4-(4-Chlorophenyl)butan-1-ol (4.16).** Prepared from **4.09** according to the general procedure for the hydrogenation of alkynols using PtO<sub>2</sub>. The crude product was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 70 mg (86%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.23 (d, *J* = 8.2 Hz, 2 H), 7.10 (d, *J* = 8.2 Hz, 2 H), 3.66 (t, *J* = 6.5 Hz, 2 H), 2.60 (t, *J* = 7.2 Hz, 2 H),

1.70-1.56 (comp, 4 H);  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  140.9, 131.7, 130.0, 128.6, 63.0, 35.2, 32.4, 27.7. IR (Neat) 3343, 2940, 2862, 2360, 1492  $\text{cm}^{-1}$ . mass spectrum (CI)  $m/z$  184.0654 [ $\text{C}_{10}\text{H}_{13}\text{OCl}$  requires 184.0655].

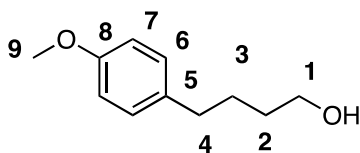
**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.23 (d,  $J = 8.2$  Hz, 2 H, C6-H), 7.10 (d,  $J = 8.2$  Hz, 2 H, C7-H), 3.66 (t,  $J = 6.5$  Hz, 2 H, C1-H), 2.60 (t,  $J = 7.2$  Hz, 2 H, C4-H), 1.70–1.56 (comp, 4 H, C2-H and C3-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  140.9 (C8), 131.7 (C5), 130.0 (C7), 128.6 (C6), 63.0 (C1), 35.2 (C4), 32.4 (C2 or C3), 27.7 (C2 or C3).



**4-(p-Tolyl)butan-1-ol (4.19).** Prepared from **4.12** according to the general procedure for the hydrogenation of the alkynols using Pd/C. The crude product was purified by silica gel chromatography eluting with  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (99:1) to yield 134 mg (99%) of the title compound as a clear oil.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.12-7.08 (comp, 4 H), 3.66 (t,  $J = 6.5$  Hz, 2 H), 2.61 (t,  $J = 7.2$  Hz, 2 H), 2.32 (s, 3 H), 1.72-1.56 (comp, 4 H);  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  139.6, 135.4, 129.3, 128.6, 62.9, 35.5, 32.5, 28.0, 21.3. IR (Neat) 3336, 2935, 2860, 1515  $\text{cm}^{-1}$ . mass spectrum (CI)  $m/z$  164.1201 [ $\text{C}_{11}\text{H}_{16}\text{O}$  requires 164.1201].

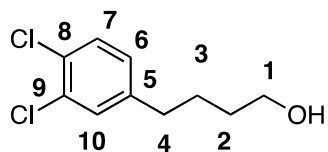
**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.12-7.08 (comp, 4 H, C6-H and C7-H), 3.66 (t,  $J = 6.5$  Hz, 2 H, C1-H), 2.61 (t,  $J = 7.2$  Hz, 2 H, C4-H), 2.32 (s, 3 H, C9-H), 1.72-1.56 (comp, 4 H, C2-H and C3-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  139.6, (C8)

135.4 (C5), 129.3 (C6), 128.6 (C7), 62.9 (C1), 35.5 (C4), 32.5 (C2 or C3), 28.0 (C2 or C3), 21.3 (C9).



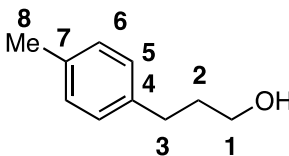
**4-(4-Methoxyphenyl)butan-1-ol (4.20).** Prepared from **4.13** according to the general procedure for the hydrogenation of alkynols using Pd/C. The crude product was purified by silica gel chromatography eluting with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (99:1) to yield 122 mg (84%) of the title compound as a clear oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.11 (d, *J* = 8.5 Hz, 2 H), 6.82 (d, *J* = 8.5 Hz, 2 H), 3.79 (s, 3 H), 3.66 (t, *J* = 6.5 Hz, 2 H), 2.59 (t, *J* = 7.5 Hz, 2 H), 1.72–1.57 (comp, 4 H); <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 157.9, 134.7, 129.5, 113.9, 63.0, 55.5, 34.9, 32.5, 28.0. IR (Neat) 3340, 2935, 2860, 1512, 1244 cm<sup>-1</sup>. mass spectrum (CI) *m/z* 180.1151 [C<sub>11</sub>H<sub>16</sub>O<sub>2</sub> requires 180.1150].

**NMR Assignments.** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.11 (d, *J* = 8.5 Hz, 2 H, C7-H), 6.82 (d, *J* = 8.5 Hz, 2 H, C6-H), 3.79 (s, 3 H, C9-H), 3.66 (t, *J* = 6.5 Hz, 2 H, C1-H), 2.59 (t, *J* = 7.5 Hz, 2 H, C4-H), 1.72–1.57 (comp, 4 H, C2-H and C3-H). <sup>13</sup>C (150 MHz, CDCl<sub>3</sub>) δ 157.9 (C8), 134.7 (C5), 129.5 (C6), 113.9 (C7), 63.0 (C1), 55.5 (C9), 34.9 (C4) 32.5 (C2 or C3), 28.0 (C2 or C3).



**4-(3,4-Dichlorophenyl)butan-1-ol (4.17).** Prepared from **4.10** according to the general procedure for the hydrogenation of alkynols using  $\text{PtO}_2$ . The crude product was purified by silica gel chromatography eluting with  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (99:1) to yield 65 mg (62%) of the title compound as a clear oil:  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.45 (d,  $J = 8.2$  Hz, 1 H), 7.27 (d,  $J = 2.1$  Hz, 1 H), 7.01 (dd,  $J = 8.2, 2.1$  Hz, 1 H), 3.66 (t,  $J = 6.2$  Hz, 2 H), 2.60 (t,  $J = 7.2$  Hz, 2 H), 1.72–1.56 (comp, 4 H);  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  142.8, 132.4, 130.6, 130.4, 129.9, 128.1, 62.8, 35.0, 32.3, 27.5. IR (Neat) 3342, 2937, 2862, 1473, 1132  $\text{cm}^{-1}$ . mass spectrum (CI)  $m/z$  218.0267 [ $\text{C}_{10}\text{H}_{12}\text{OCl}_2$  requires 218.0265].

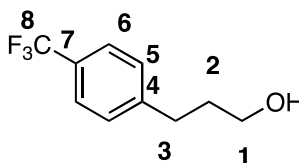
**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.45 (d,  $J = 8.2$  Hz, 1 H, C7-H), 7.27 (d,  $J = 2.1$  Hz, 1 H, C10-H), 7.01 (dd,  $J = 8.2, 2.1$  Hz, 1 H, C6-H), 3.66 (t,  $J = 6.2$  Hz, 2 H, C1-H), 2.60 (t,  $J = 7.2$  Hz, 2 H, C4-H), 1.72–1.56 (comp, 4 H, C2-H and C3-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  142.8 (C5), 132.4 (C9), 130.6 (C8), 129.9 (C10), 128.1 (C7), 62.8 (C1), 35.0 (C4), 32.3 (C2 or C3), 27.5 (C2 or C3).



**3-(p-Tolyl)propan-1-ol (4.22).** Prepared from **4.15** according to the general procedure for the hydrogenation of alkynols using  $\text{Pd/C}$ . The crude product was purified by silica gel chromatography eluting with  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (99:1) to yield 70 mg (99%) of the title compound as a clear oil.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.18–7.16 (comp, 4 H),

3.63 (t,  $J = 6.5$  Hz, 2 H), 2.62 (t,  $J = 7.2$  Hz, 2 H), 2.35 (s, 3 H), 1.72-1.56 (comp, 2 H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  139.0, 135.6, 129.3, 128.6, 62.5, 34.6, 31.9, 21.3. IR (Neat) 3336, 2935, 2860, 1515  $\text{cm}^{-1}$ .

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.18-7.16 (comp, 4 H, C5-H and C6-H), 3.63 (t,  $J = 6.5$  Hz, 2 H, C1-H), 2.62 (t,  $J = 7.2$  Hz, 2 H, C3-H), 2.35 (s, 3 H, C8-H), 1.72-1.56 (comp, 2 H, C2-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  139.0 (C4), 135.6 (C7), 129.3 (C5), 128.6 (C6), 62.5 (C1), 34.6 (C3), 31.9 (C2), 21.3 (C8).



**3-(4-(Trifluoromethyl)phenyl)propan-1-ol (4.21).** Prepared from **4.14** according to the general procedure for the hydrogenation of alkynols using Pd/C. The crude product was purified by silica gel chromatography eluting with  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (99:1) to yield 126 mg (57%) of the title compound as a clear oil.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.52 (d,  $J = 7.8$  Hz, 2 H), 7.27 (d,  $J = 7.8$  Hz, 2 H), 3.69 (t,  $J = 6.2$  Hz, 2 H), 2.76 (t,  $J = 7.9$  Hz, 2 H), 1.93 (comp, 2 H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  146.2, 129.0, 128.3 (q,  $J = 32.3$  Hz), 125.6, 123.2, 62.2, 34.1, 32.1. IR (Neat) 3344, 2942, 2870, 1327, 1120  $\text{cm}^{-1}$ . mass spectrum (ESI)  $m/z$  218.0915 [ $\text{C}_{10}\text{H}_{11}\text{OF}_3$  requires 218.0918].

**NMR Assignments.**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.52 (d,  $J = 7.8$  Hz, 2 H, C5-H), 7.27 (d,  $J = 7.8$  Hz, 2 H, C6-H), 3.69 (t,  $J = 6.2$  Hz, 2 H, C1-H), 2.76 (t,  $J = 7.9$  Hz, 2 H, C4-H), 1.93 (comp, 2 H, C2-H).  $^{13}\text{C}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  146.2 (C6), 129.0 (C5), 128.3 (q,  $J = 32.3$  Hz, C7), 125.6 (C8), 123.2 (C4), 62.2 (C1), 34.1 (C3), 32.1 (C2).

### 5.3 General Procedure for Performing ITC

Isothermal titration calorimetry (ITC) experiments were performed on a VP-ITC (GE Healthcare) at 298.15 K in the appropriate buffer at the given pH 7.4. A minimum of two experiments with each ligand were performed. Protein and ligand solutions were filtered using sterile 0.2  $\mu$ M EO minisart filters (Sartorius Stedim Biotech). Protein concentration was determined by UV spectroscopy (for MUP-I:  $\epsilon_{280} = 12,630 \text{ M}^{-1}\text{cm}^{-1}$ , for Grb2 SH2:  $\epsilon_{280} = 15,000 \text{ M}^{-1}\text{cm}^{-1}$ ). Solutions were stored at 4 °C until needed and degassed for 15 m. at 23 °C using a ThermoVac vacuum chamber (GE Healthcare) immediately prior to ITC experiments.

The dialysate was re-filtered using a 0.2  $\mu$ M cellulose membrane filter (Whatman) and thoroughly degassed before being used to make ligand solutions. Ligand solutions were prepared from degassed PBS dialysate that was re-filtered using a 0.2  $\mu$ M cellulose membrane filter (Whatman).

Titration experiments consisted of an initial 2  $\mu$ L injection followed by 24 x 10  $\mu$ L injections spaced 240 s. apart using a stirring speed of 300 rpm. The cell was cleaned prior to each run with 1 M NaOH followed by extensive rinsing with both water and the dialysate. Between runs of different ligands the syringe was flushed extensively with the latter two solutions, but only PBS dialysate was used between titrations with the same ligand.

#### *Analysis of ITC Data*

ITC data were analyzed in Origin 5.0 (MicroCal). Blank experiments (ligand into buffer) yielded consistent heats of dilution and were fit to a line with zero slope to obtain the average or blank value. Following subtraction of the blank values and removal of

data corresponding to the initial 2  $\mu$ L injection, the adjusted data were fit to a one-site model of the Wiseman isotherm.

#### *Treatment of Experimental Error*

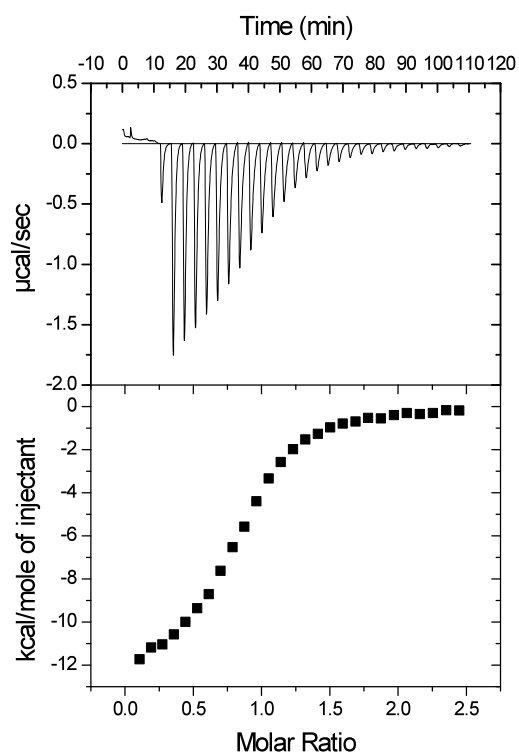
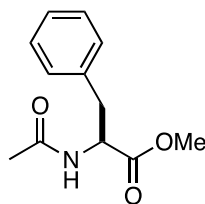
Error in the observed free-energies and enthalpies of binding was defined as the standard error in the mean of replicate experiments. These errors were propagated through the Gibbs equation to give errors in the observed entropies.

#### *Determination of Ligand Molecular Surface Area*

The ligands in the X-ray crystal structures were removed from the complexes and their Connolly or molecular surface areas were calculated based upon their bound conformations using Macromodel v9.1.



## 5.4 Average Thermodynamic Parameters and Representative ITC traces



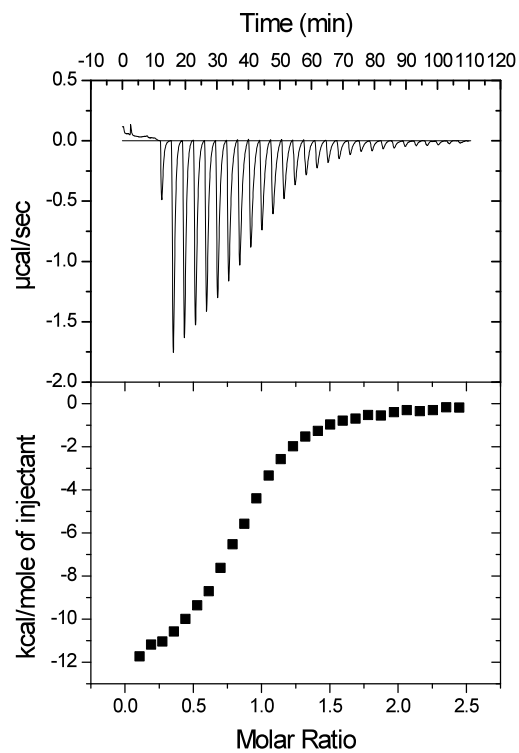
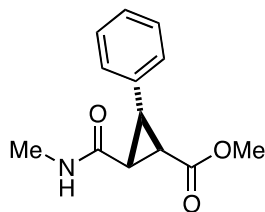
$$T = 298 \text{ K}$$

$$K_a = 2.8 \times 10^5$$

$$\Delta G^\circ = -7.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -11.3 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 3.6 \text{ kcal} \cdot \text{mol}^{-1}$$



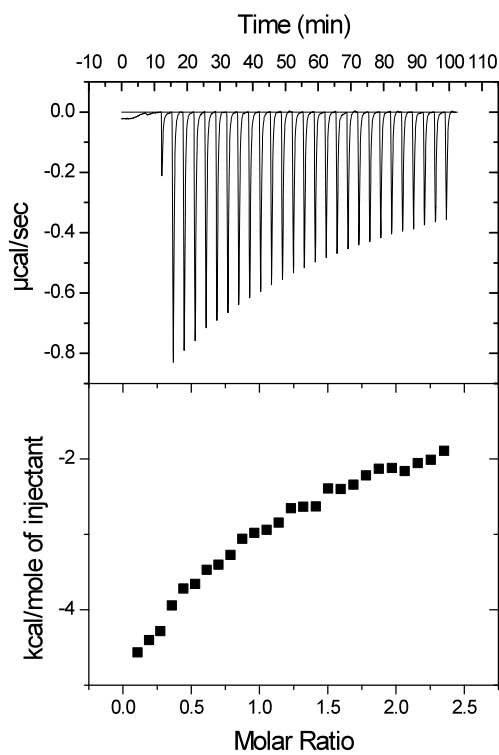
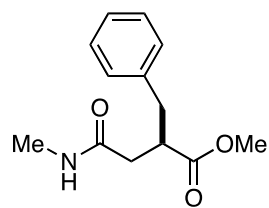
$$T = 298 \text{ K}$$

$$K_a = 2.4 \times 10^5 \text{ M}^{-1}$$

$$\Delta G^\circ = -7.2 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -12.4 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 5.1 \text{ kcal}\cdot\text{mol}^{-1}$$



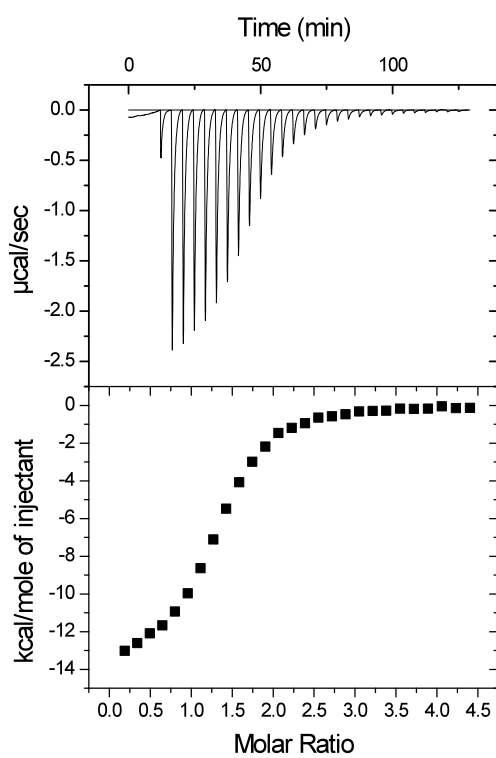
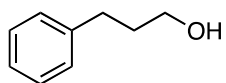
$$T = 298 \text{ K}$$

$$K_a < 1 \times 10^3 \text{ M}^{-1}$$

$$\Delta G^\circ = \text{na}$$

$$\Delta H^\circ = \text{na}$$

$$-T\Delta S^\circ = \text{na}$$



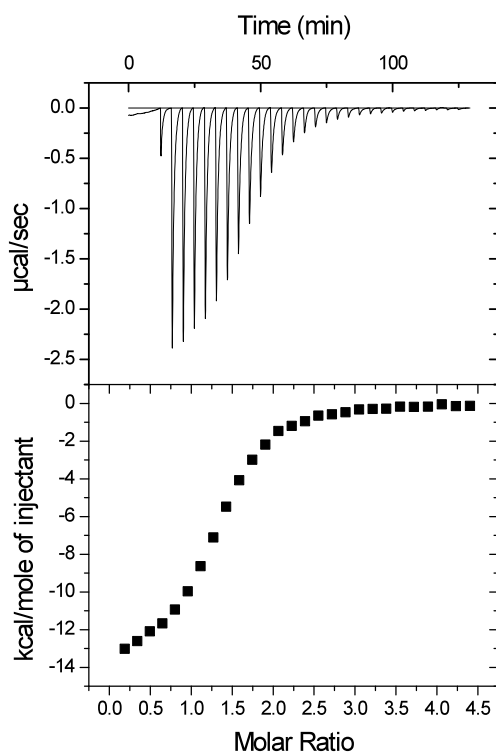
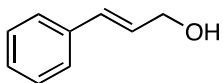
$$T = 298 \text{ K}$$

$$K_a = 1 \times 10^3 \text{ M}^{-1}$$

$$\Delta G^\circ = -6.9 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -12.2 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 5.3 \text{ kcal}\cdot\text{mol}^{-1}$$



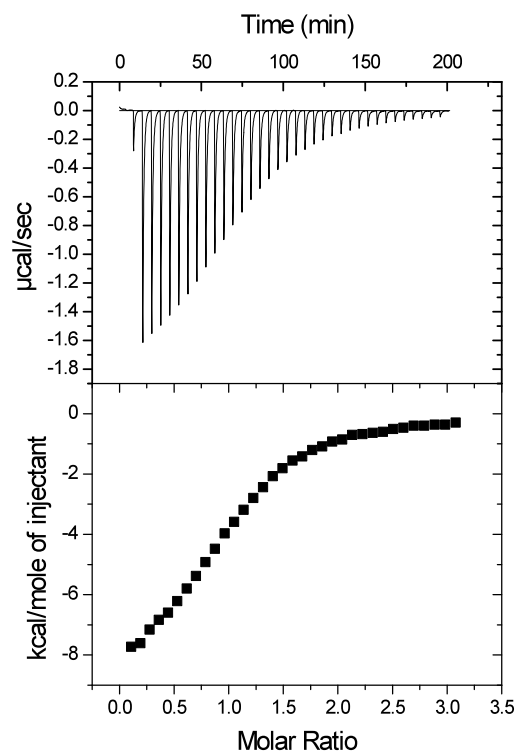
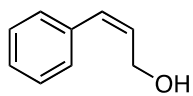
$$T = 298 \text{ K}$$

$$K_a < 1 \times 10^3 \text{ M}^{-1}$$

$$\Delta G^\circ = -6.9 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -11.4 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 4.6 \text{ kcal}\cdot\text{mol}^{-1}$$



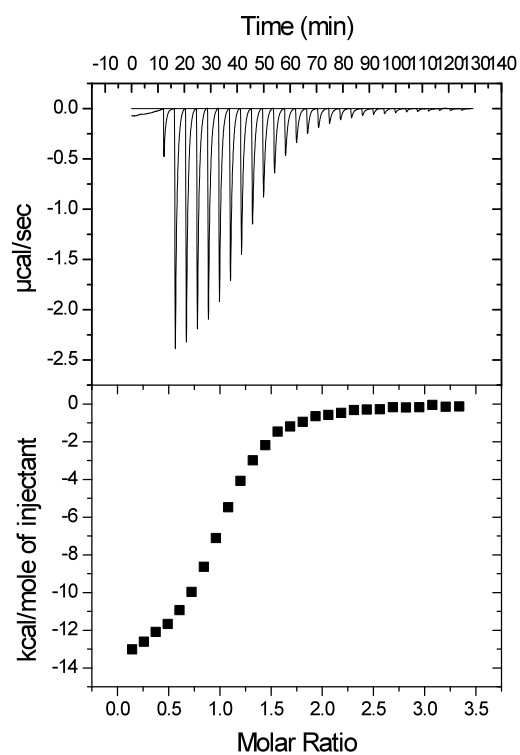
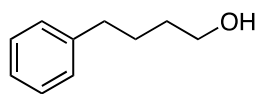
$$T = 298 \text{ K}$$

$$K_a = 8.2 \times 10^4 \text{ M}^{-1}$$

$$\Delta G^\circ = -6.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.7 \text{ kcal} \cdot \text{mol}^{-1}$$



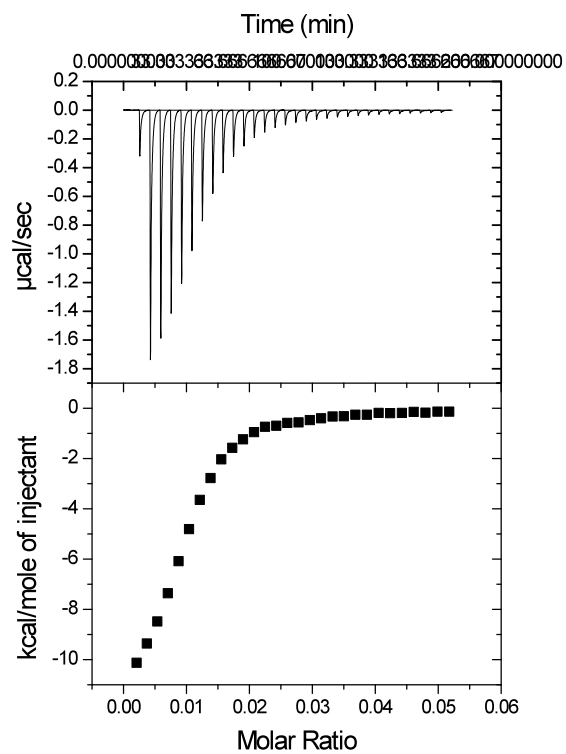
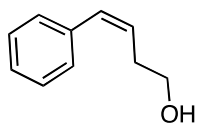
$$T = 298 \text{ K}$$

$$K_a = 3.0 \times 10^5 \text{ M}^{-1}$$

$$\Delta G^\circ = -7.5 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -13.8 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 6.4 \text{ kcal}\cdot\text{mol}^{-1}$$



$$T = 298 \text{ K}$$

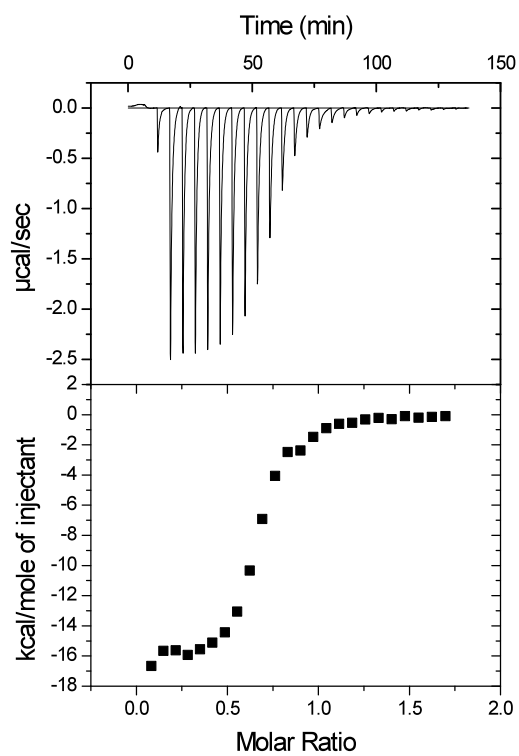
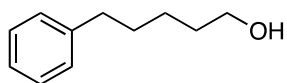
$$K_a = 3.0 \times 10^5 \text{ M}^{-1}$$

$$\Delta G^\circ = -7.3 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -11.0 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 3.7 \text{ kcal}\cdot\text{mol}^{-1}$$





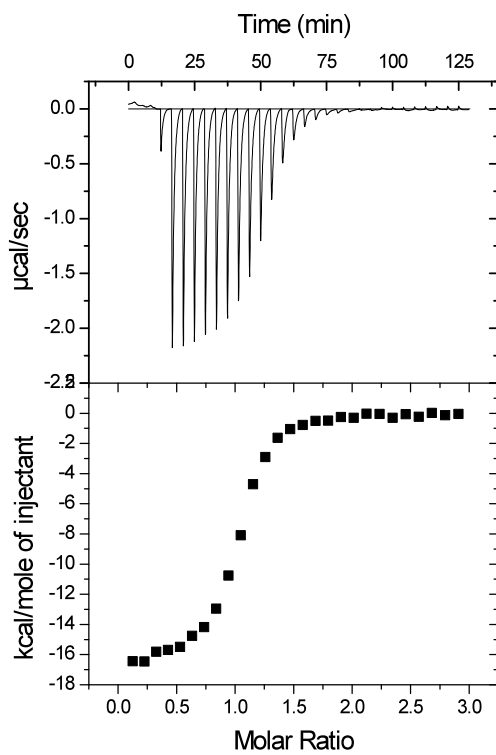
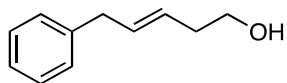
$$T = 298 \text{ K}$$

$$K_a = 1.1 \times 10^6 \text{ M}^{-1}$$

$$\Delta G^\circ = -8.3 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -16.5 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 8.2 \text{ kcal}\cdot\text{mol}^{-1}$$



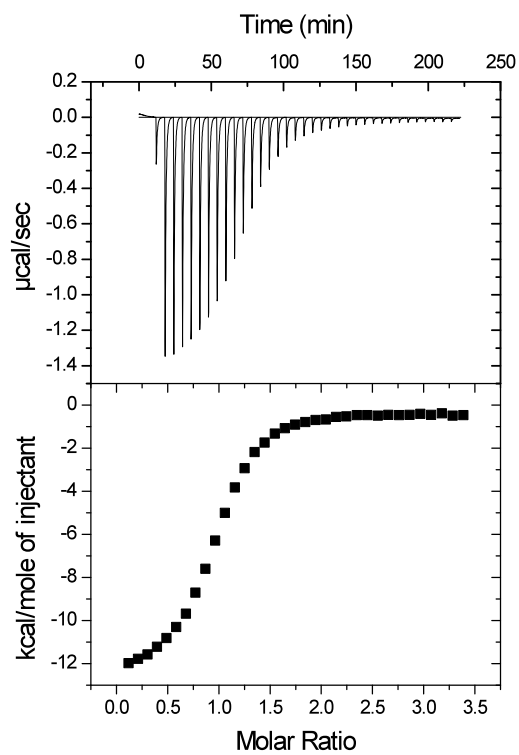
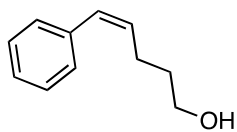
$$T = 298 \text{ K}$$

$$K_a < 9.9 \times 10^5 \text{ M}^{-1}$$

$$\Delta G^\circ = -7.2 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -12.7 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 4.8 \text{ kcal}\cdot\text{mol}^{-1}$$



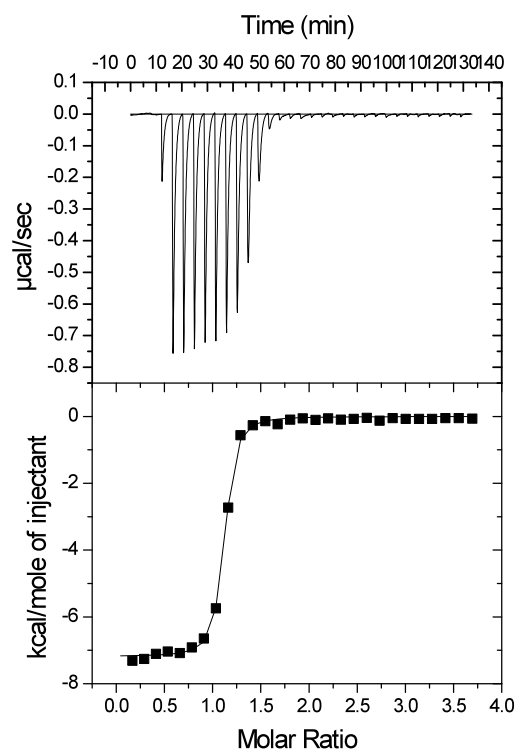
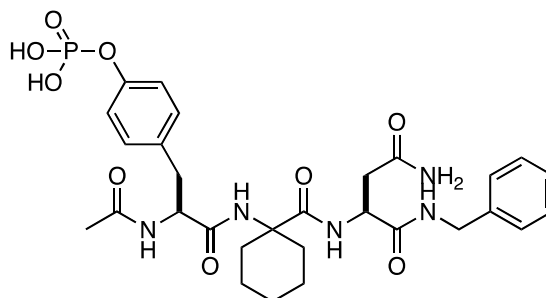
$$T = 298 \text{ K}$$

$$K_a = 4.3 \times 10^5 \text{ M}^{-1}$$

$$\Delta G^\circ = -7.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -13.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 5.4 \text{ kcal} \cdot \text{mol}^{-1}$$



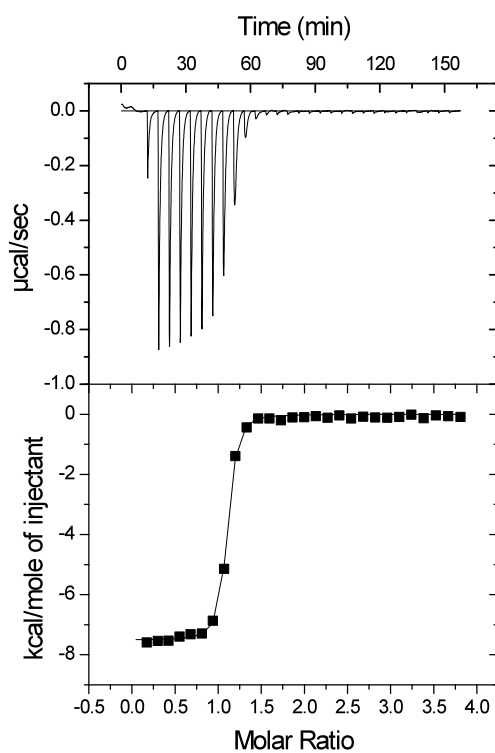
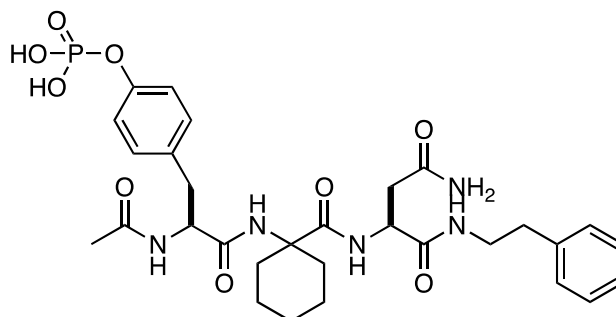
$$T = 298 \text{ K}$$

$$K_a = 1.5 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -7.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 2.1 \text{ kcal} \cdot \text{mol}^{-1}$$



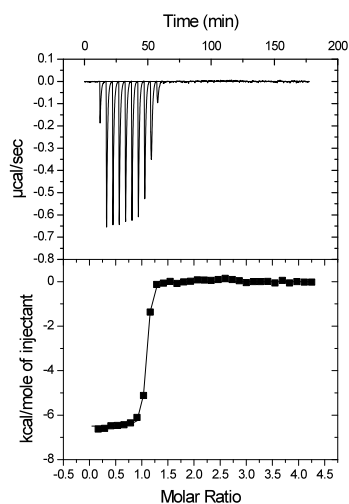
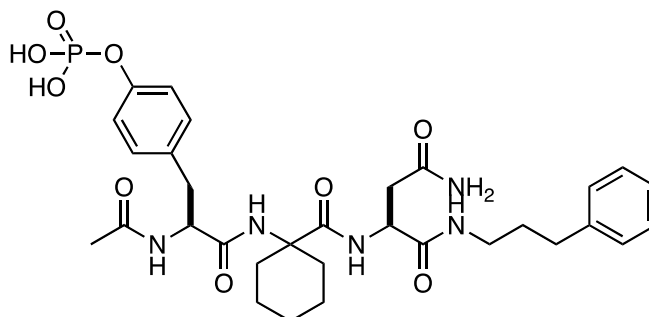
$$T = 298 \text{ K}$$

$$K_a = 2.6 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -7.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 2.1 \text{ kcal} \cdot \text{mol}^{-1}$$



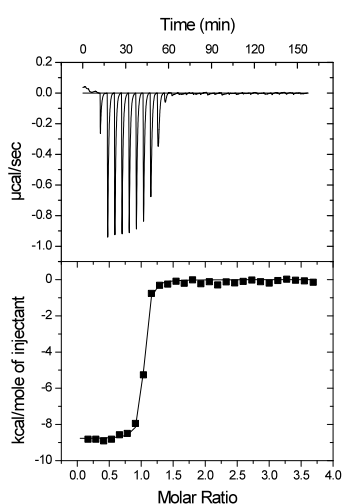
T = 288 K

$$K_a = 2.6 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -6.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -3.4 \text{ kcal} \cdot \text{mol}^{-1}$$



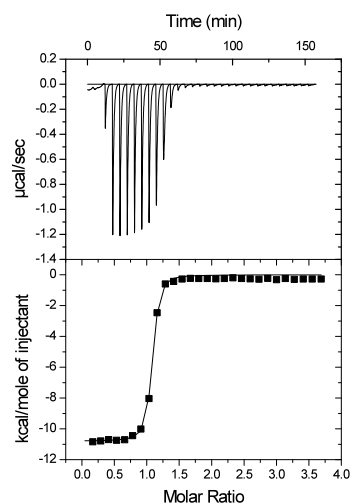
T = 298 K

$$K_a = 2.6 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.4 \text{ kcal} \cdot \text{mol}^{-1}$$



T = 308 K

$$K_a = 2.6 \times 10^7 \text{ M}^{-1}$$

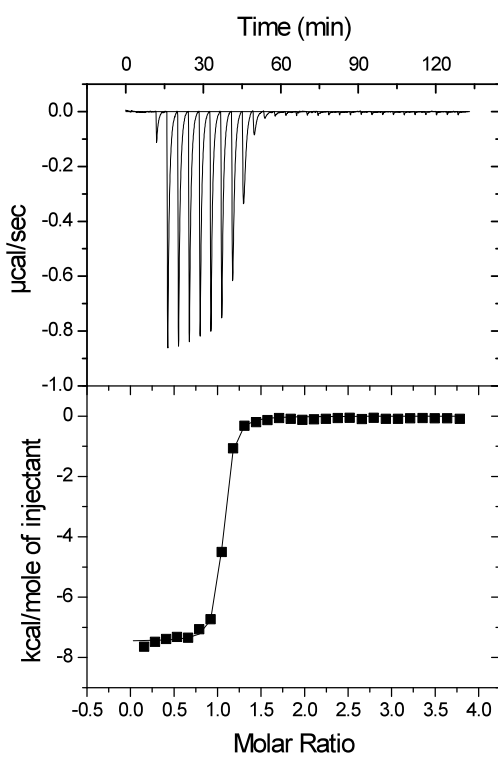
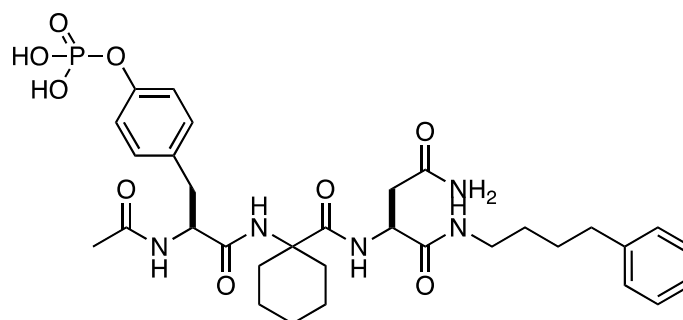
$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -10.5 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -0.3 \text{ kcal} \cdot \text{mol}^{-1}$$

1

1



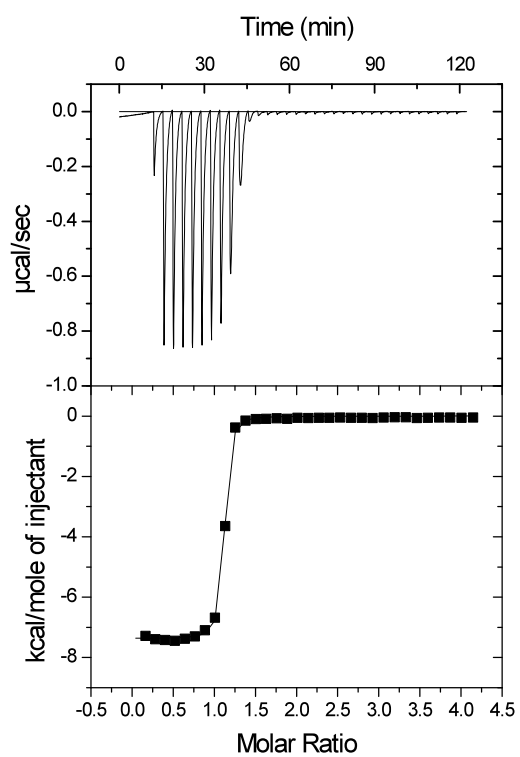
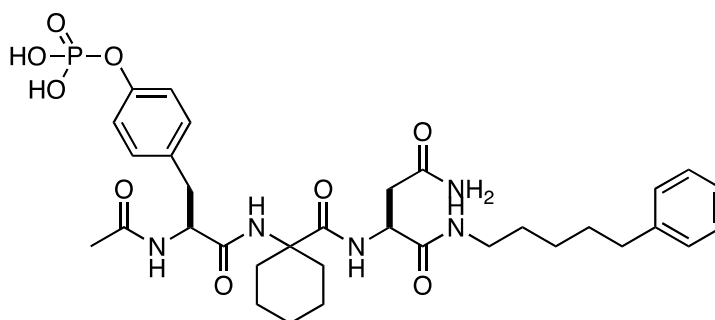
$$T = 298 \text{ K}$$

$$K_a = 4.2 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -7.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -2.0 \text{ kcal} \cdot \text{mol}^{-1}$$



$$T = 298 \text{ K}$$

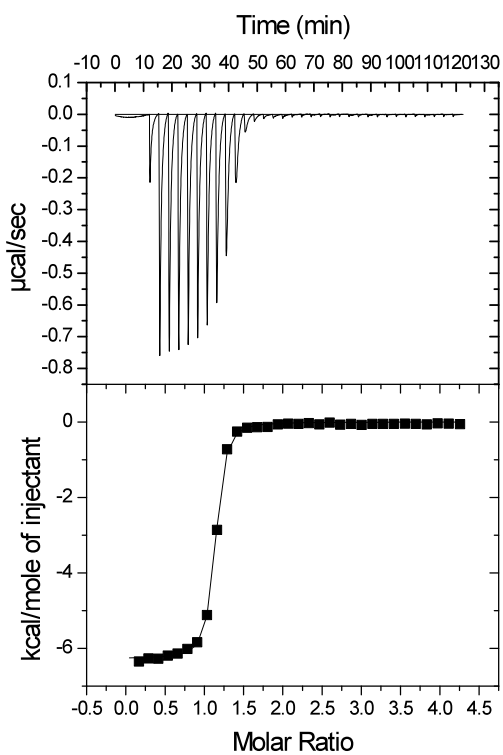
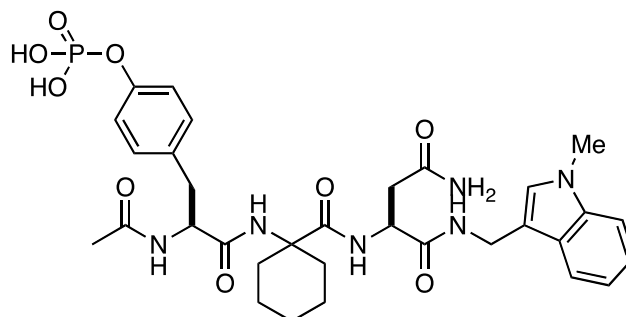
$$K_a = 2.8 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.2 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -7.9 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = -2.3 \text{ kcal}\cdot\text{mol}^{-1}$$





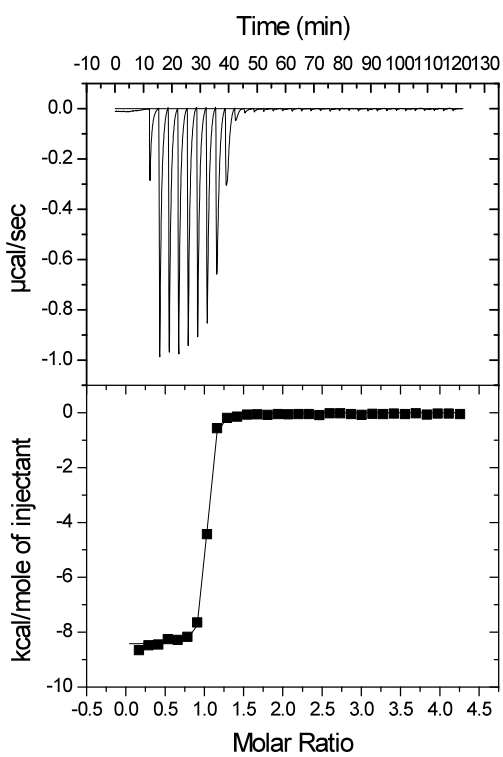
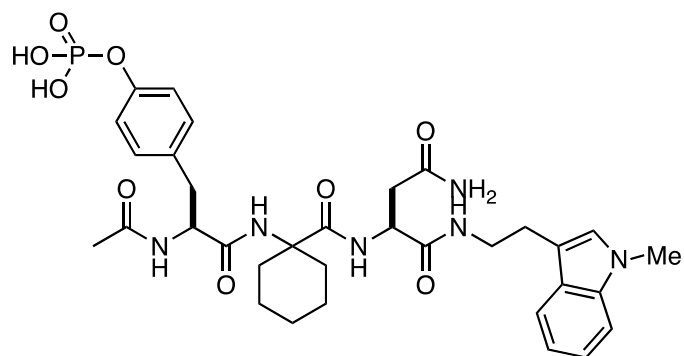
$$T = 298 \text{ K}$$

$$K_a = 9.6 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -6.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -2.9 \text{ kcal} \cdot \text{mol}^{-1}$$



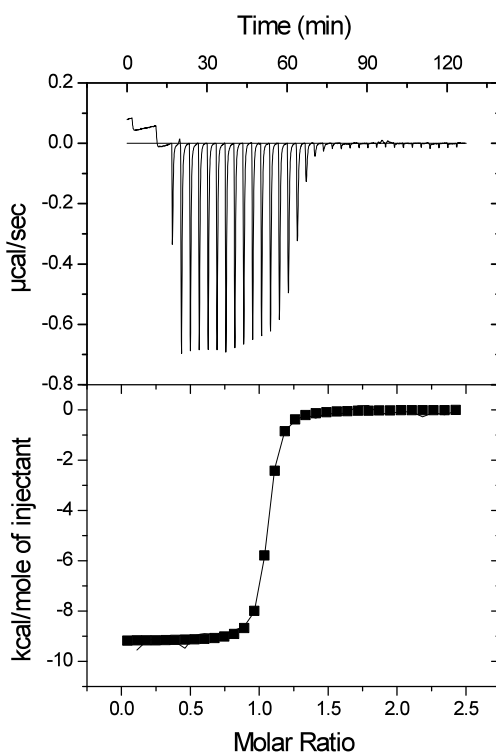
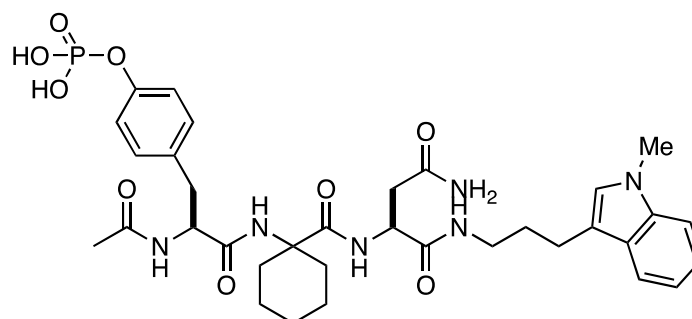
$$T = 298 \text{ K}$$

$$K_a = 2.4 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.1 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.1 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -2.0 \text{ kcal} \cdot \text{mol}^{-1}$$



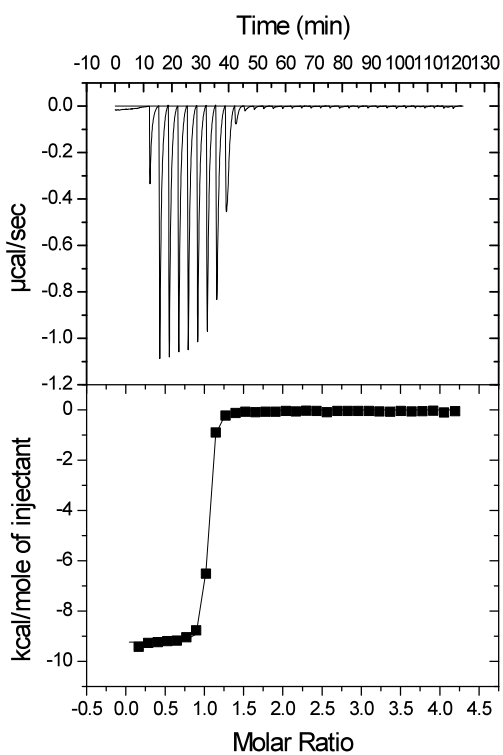
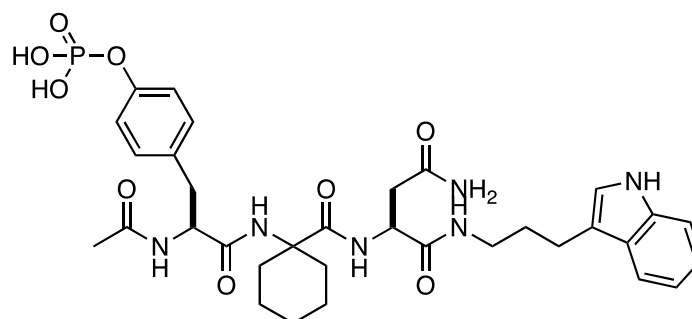
$$T = 298 \text{ K}$$

$$K_a = 4.9 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -19.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -0.2 \text{ kcal} \cdot \text{mol}^{-1}$$



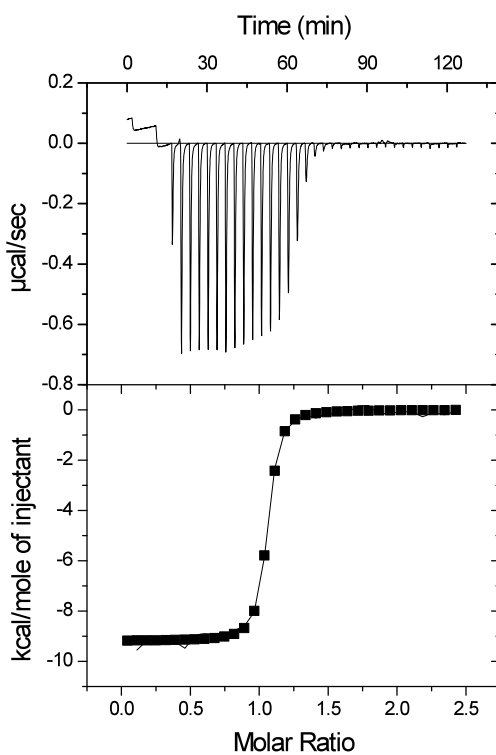
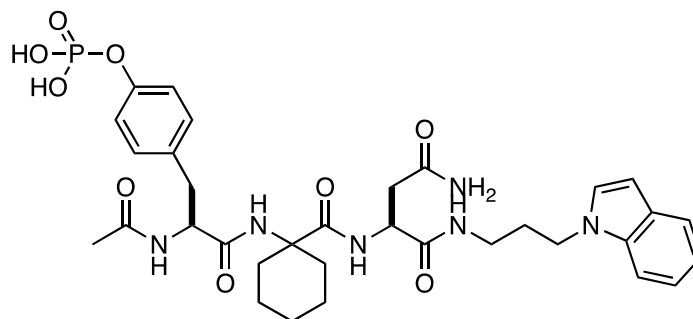
$$T = 298 \text{ K}$$

$$K_a = 2.8 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.3 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -9.1 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.2 \text{ kcal} \cdot \text{mol}^{-1}$$



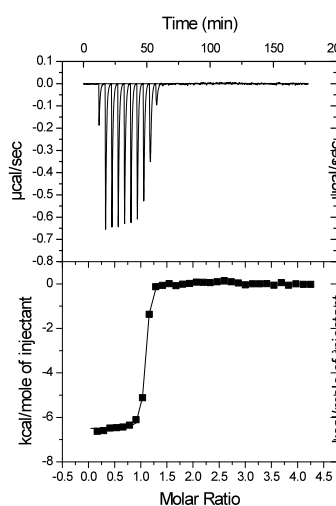
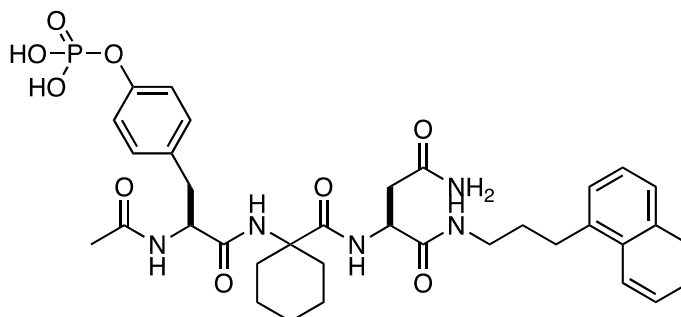
$$T = 298 \text{ K}$$

$$K_a = 2.8 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -9.1 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.1 \text{ kcal} \cdot \text{mol}^{-1}$$



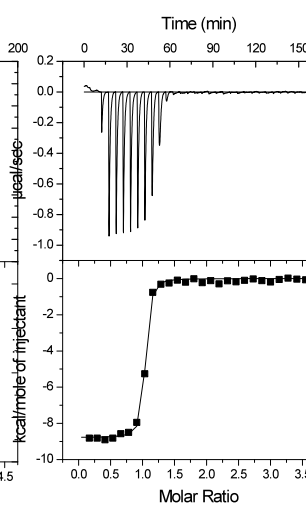
T = 288 K

$$K_a = 5.1 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -6.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -3.4 \text{ kcal} \cdot \text{mol}^{-1}$$



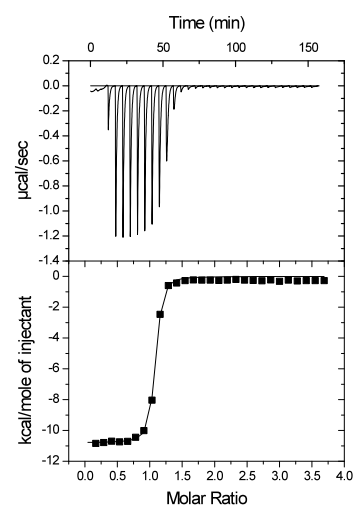
T = 298 K

$$K_a = 6.0 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.6 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -9.6 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.2 \text{ kcal} \cdot \text{mol}^{-1}$$



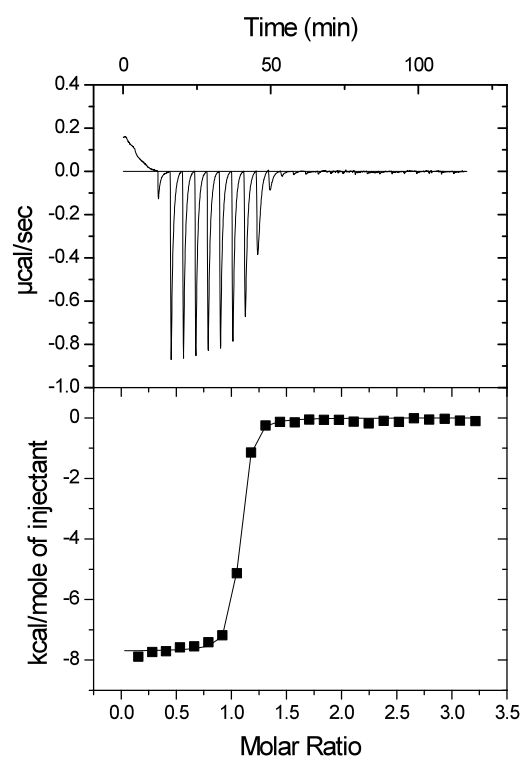
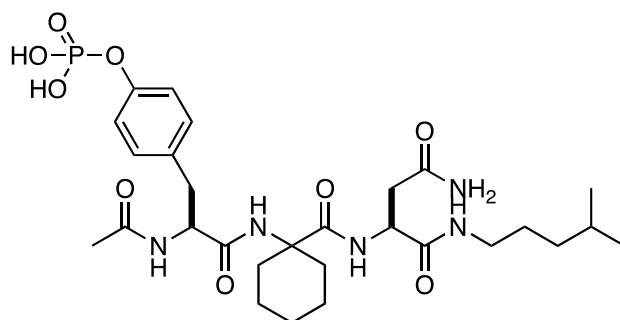
T = 308 K

$$K_a = 2.6 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -11.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 3.1 \text{ kcal} \cdot \text{mol}^{-1}$$



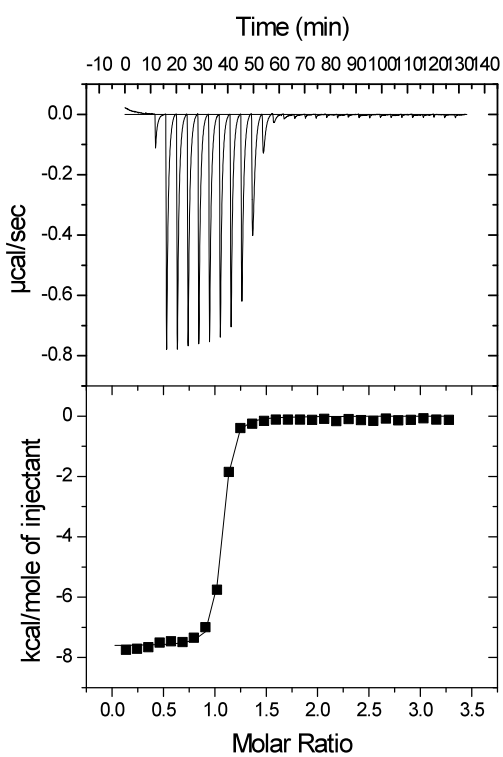
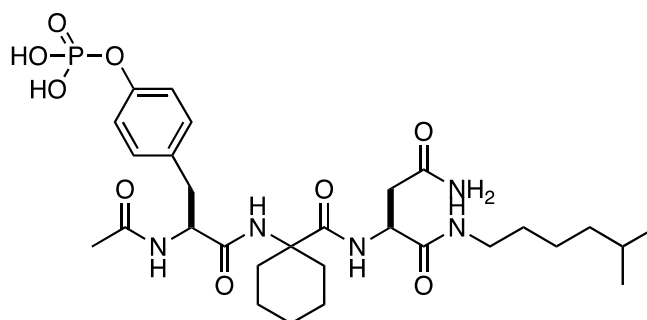
$$T = 298 \text{ K}$$

$$K_a = 1.7 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -7.8 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = -2.1 \text{ kcal}\cdot\text{mol}^{-1}$$



$$T = 298 \text{ K}$$

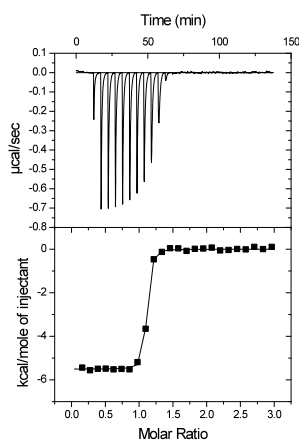
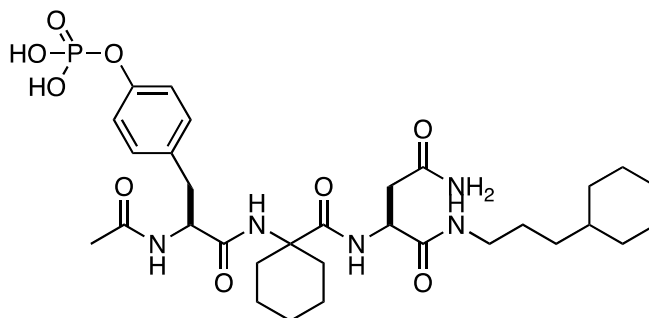
$$K_a = 1.7 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -7.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -2.1 \text{ kcal} \cdot \text{mol}^{-1}$$





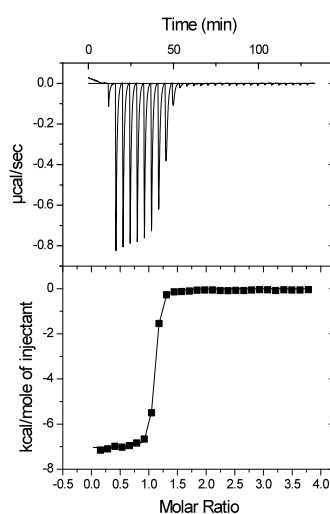
T = 298 K

$$K_a = 4.1 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -5.5 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -4.7 \text{ kcal} \cdot \text{mol}^{-1}$$



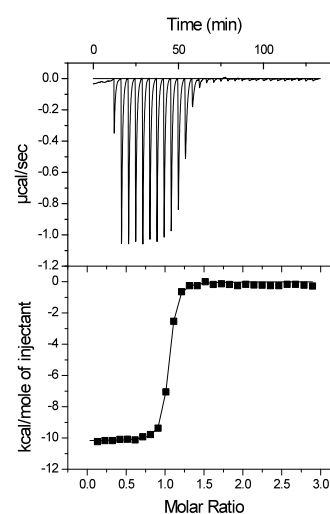
T = 298 K

$$K_a = 2.8 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -9.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.0 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.9 \text{ kcal} \cdot \text{mol}^{-1}$$



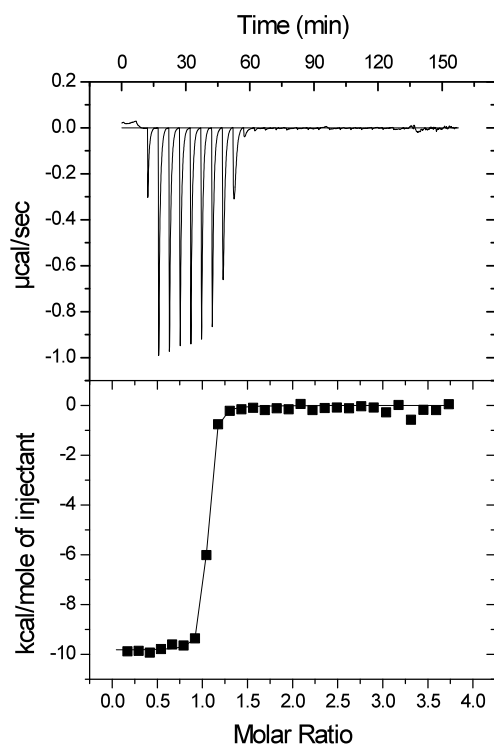
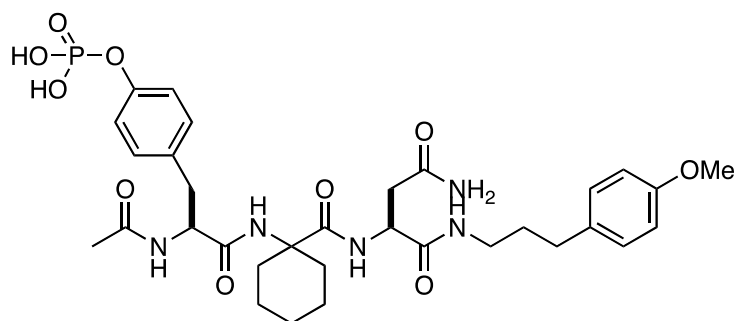
T = 208 K

$$K_a = 2.8 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.2 \text{ kcal} \cdot \text{mol}^{-1}$$



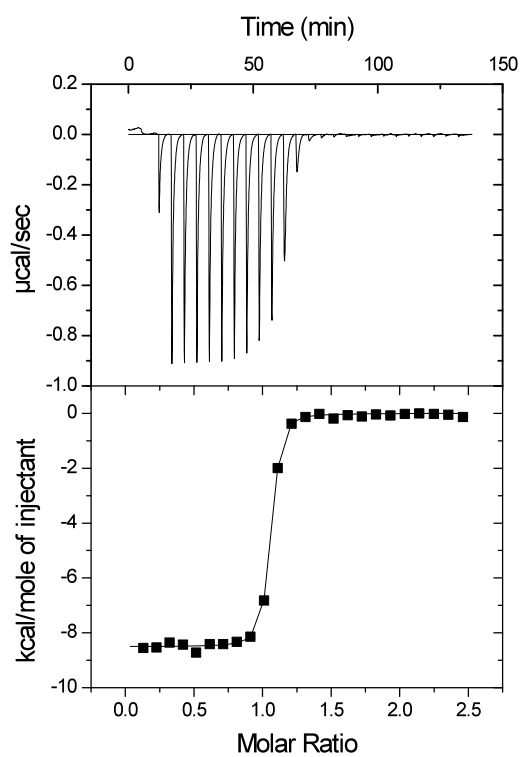
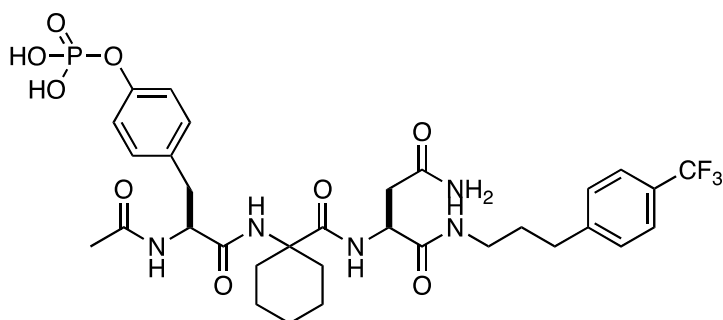
$$T = 298 \text{ K}$$

$$K_a = 4.7 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.3 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -9.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.2 \text{ kcal} \cdot \text{mol}^{-1}$$



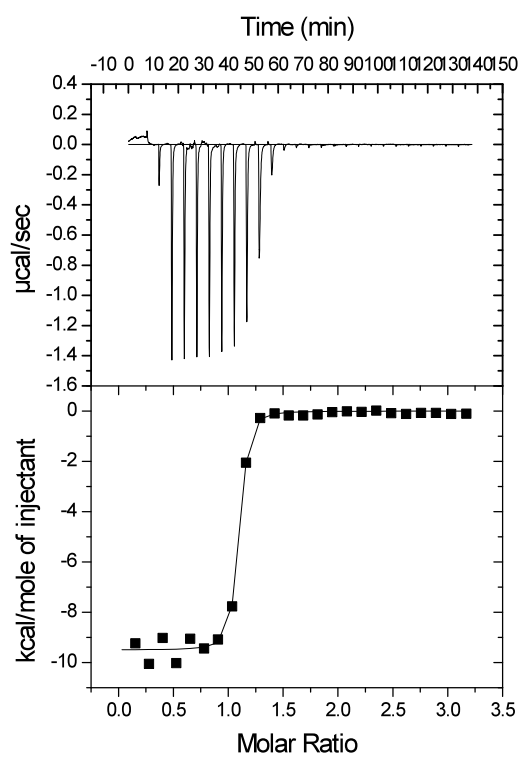
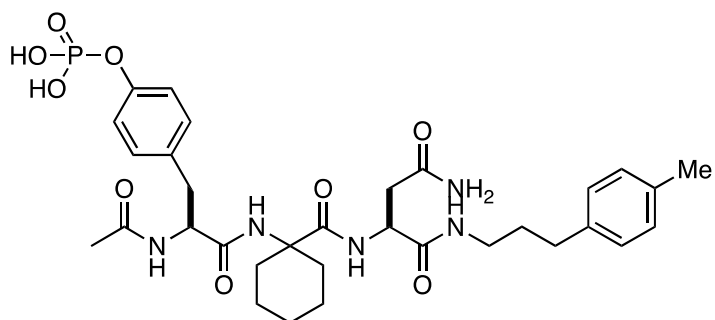
$$T = 298 \text{ K}$$

$$K_a = 3.0 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.1 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.7 \text{ kcal} \cdot \text{mol}^{-1}$$



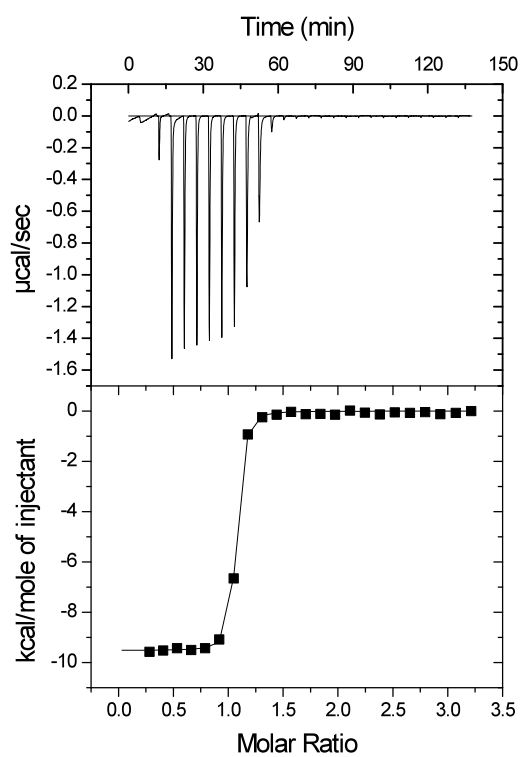
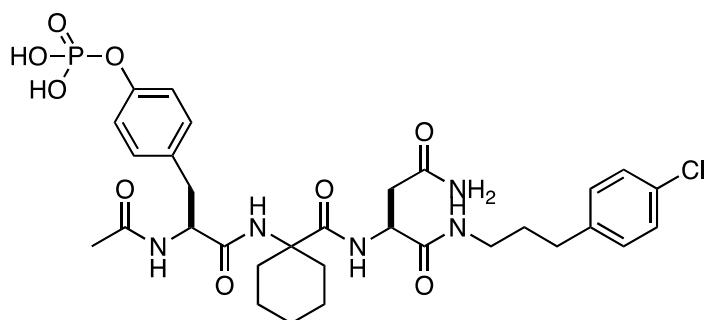
$$T = 298 \text{ K}$$

$$K_a = 4.7 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.3 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -8.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.7 \text{ kcal} \cdot \text{mol}^{-1}$$



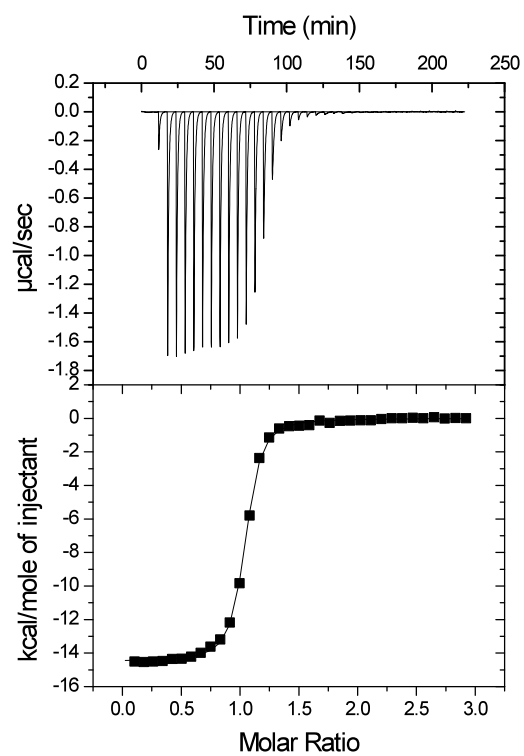
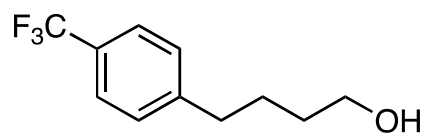
$$T = 298 \text{ K}$$

$$K_a = 5.1 \times 10^7 \text{ M}^{-1}$$

$$\Delta G^\circ = -10.5 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -9.5 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = -1.0 \text{ kcal} \cdot \text{mol}^{-1}$$



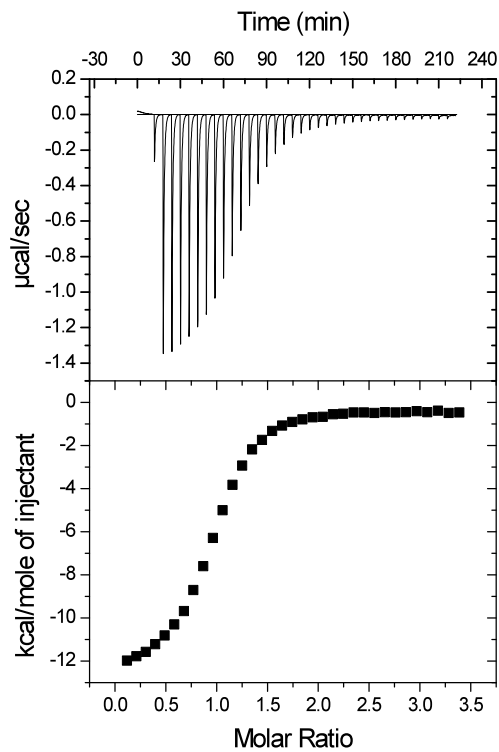
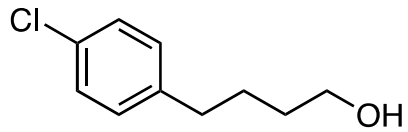
$$T = 298 \text{ K}$$

$$K_a = 2.8 \times 10^6 \text{ M}^{-1}$$

$$\Delta G^\circ = -8.8 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -14.0 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 5.2 \text{ kcal} \cdot \text{mol}^{-1}$$



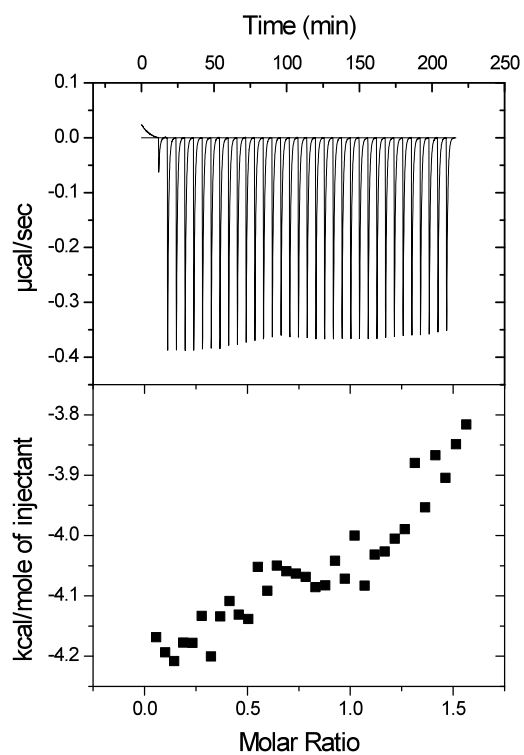
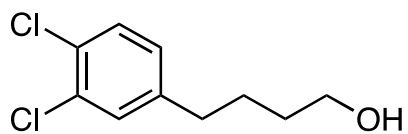
$$T = 298 \text{ K}$$

$$K_a = 1.6 \times 10^6 \text{ M}^{-1}$$

$$\Delta G^\circ = -8.4 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta H^\circ = -12.2 \text{ kcal}\cdot\text{mol}^{-1}$$

$$-T\Delta S^\circ = 3.8 \text{ kcal}\cdot\text{mol}^{-1}$$



$$T = 298 \text{ K}$$

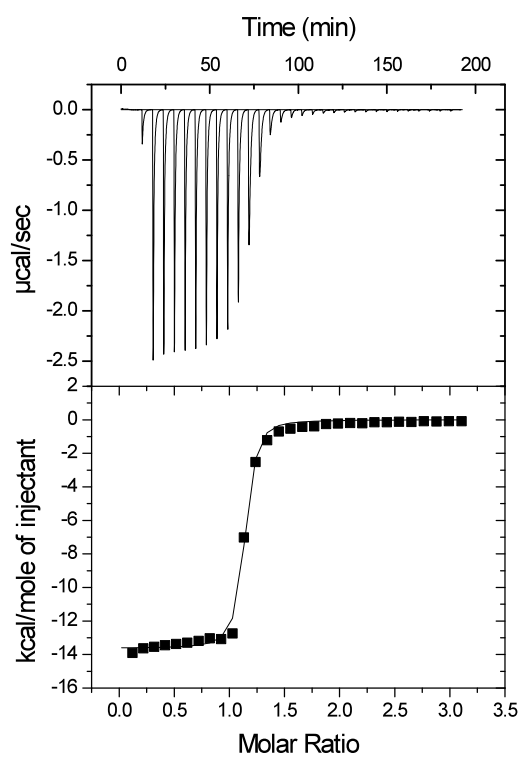
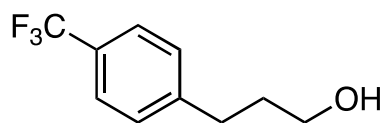
$$K_a < 1 \times 10^3 \text{ M}^{-1}$$

$$\Delta G^\circ = \text{na}$$

$$\Delta H^\circ = \text{na}$$

$$-T\Delta S^\circ = \text{na}$$





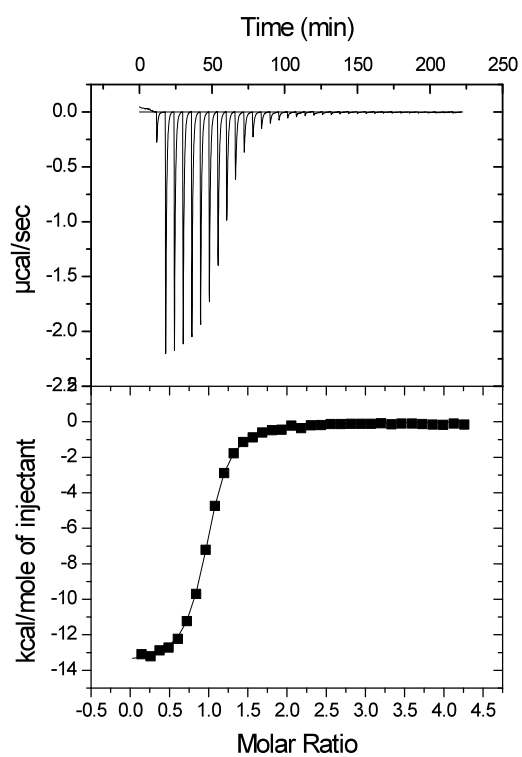
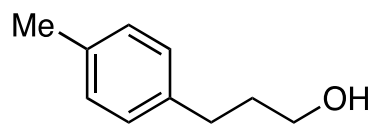
$$T = 298 \text{ K}$$

$$K_a = 3.4 \times 10^6 \text{ M}^{-1}$$

$$\Delta G^\circ = -8.9 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -15.3 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 6.4 \text{ kcal} \cdot \text{mol}^{-1}$$



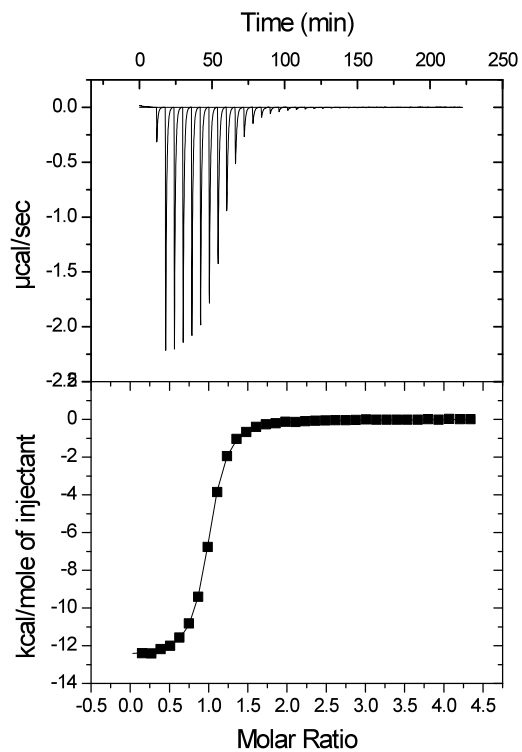
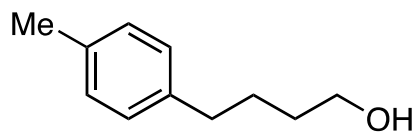
$$T = 298 \text{ K}$$

$$K_a = 8.8 \times 10^5 \text{ M}^{-1}$$

$$\Delta G^\circ = -8.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -13.7 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 5.6 \text{ kcal} \cdot \text{mol}^{-1}$$



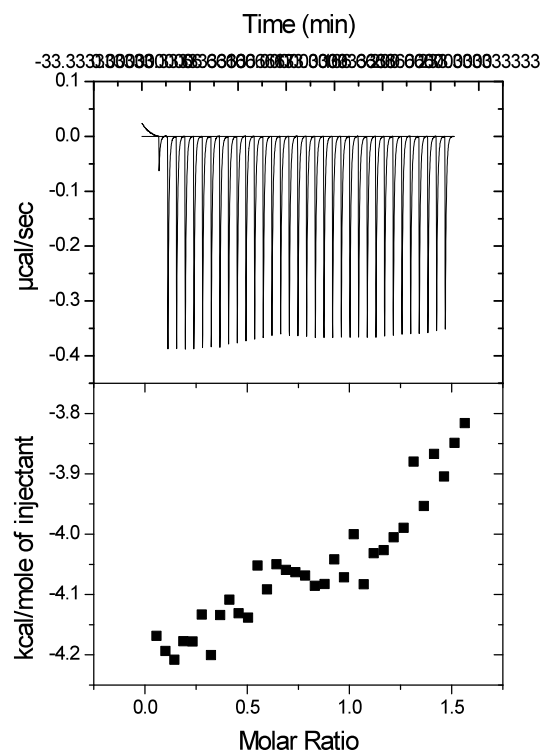
$$T = 298 \text{ K}$$

$$K_a = 1.6 \times 10^6 \text{ M}^{-1}$$

$$\Delta G^\circ = -8.2 \text{ kcal} \cdot \text{mol}^{-1}$$

$$\Delta H^\circ = -11.4 \text{ kcal} \cdot \text{mol}^{-1}$$

$$-T\Delta S^\circ = 5.2 \text{ kcal} \cdot \text{mol}^{-1}$$


$$K_a < 1 \times 10^3 \text{ M}^{-1}$$

$$\Delta H^{\circ} = na$$

## References

- <sup>i</sup> For reports on recent scoring functions, see: (a) Raha, K.; Merz, K. M. *J. Med. Chem.* **2005**, *48*, 4558-4575. (b.) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Bioph. Biom.* **2007**, *36*, 21-42. (c) Gao, C.; Park, M. S.; Stern, H. A. *Biophys. J.* **2010**, *98*, 901-910. (d) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc Natl Acad Sci USA*, **2012**, *109*, 1937-1942.
- <sup>ii</sup> For reports on limitations of scoring functions, see: (a) Warren, G.L.; Andrews, C. W.; Capello, A.M.; Clarke, B.; LaLonde, J.; Lambert, M.H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Semger, S.; Tedesco, G.; Wall, I.D.; Woolven, J.M.; Peishoff, C.E.; Head, M.S. *J. Med. Chem.* **2006**, *11*, 580-594. (b) Klebe, G.; *Drug Discovery Today* **2006**, *11*, 580-594. (c) Damn, K.L.; Carlson, H. A. *J. Am. Chem. Soc.* **2007**, *129*, 8225-8235.
- <sup>iii</sup> Beggs, M. "HTS: where next?" *Drug Discov. World* **2000**, *1*, 25-27.
- <sup>iv</sup> Mander, T. "Beyond  $\mu$ HTS: Ridiculously HTS?" *Drug Discov. Today*, **2000**, *5*, 223-225.
- <sup>v</sup> Ruiz-Garcia, A.; Barnejo, M.; Moss, A.; Casabo, V.G. "Pharmacokinetics in Drug Discovery" *J. Pharm. Sci.*, **2008**, 654-690.
- <sup>vi</sup> Kolb, A.; Neumann, K., Mathis, G. "New Developments in HTS Screening Technology". *Pharm. Manuf. Int.* **1996**, *4*, 31-39.
- <sup>vii</sup> Pope, A.J; Haupts, U.M.; Moore, K.L. Homogeneous Fluorescence Readouts for Minuturized High Throughput Screening: Theory and Practice. *Drug Discovery Today*, **1999**, *4*, 350-362.
- <sup>viii</sup> Figure taken from ref 3, used by permission.
- <sup>ix</sup> *Industrialization of Drug Discovery: From target Selection through Lead Optimization*; Handen, J. S., Ed.; Drug Discovery Series; CRC Press: Boca Raton, FL, 2005.
- <sup>x</sup> Martin, S.F. "Preorganization in biological systems: are conformational constraints worth the energy?" *Pure Appl. Chem.* **2007**, *79*, 193-200.
- <sup>xi</sup> Ruben, A. J.; Kiso, Y.; Freire, E. "Overcoming Roadblocks in Lead Optimization: A Thermodynamic Prospective" *Chem. Biol. Drug. Des.* **2006**, *67*, 2-4.

- <sup>xii</sup> Homans, S. W. "Water, water everywhere—except where it matters?" *Drug Discovery Today* **2007**, *12*, 534-539.
- <sup>xiii</sup> Olsson, T. S.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. "The Thermodynamics of Protein–Ligand Interaction and Solvation: Insights for Ligand Design" *J. Mol. Biol.* **2008**, *384*, 1002-1017.
- <sup>xiv</sup> Homans, S.W.; *Bioactive Conformation I*, Springer: Berlin, 2007, 272.
- <sup>xv</sup> Malham, R.; Johnstone, S.; Bingham, R.J.; Barratt, E.; Phillips, S. E. V.; Laughton, C. A.; Homans, S. W. "Strong Solute-Solute Dispersive Interactions in a Protein-Ligand Complex" *J. Am. Chem. Soc.* **2005**, *437*, 640-647.
- <sup>xvi</sup> Cheng, Y.-C.; Prusoff, W. H. "Relationship Between the Inhibition Constant (K<sub>i</sub>) and the Concentration of Inhibitor Which Causes 50 per cent Inhibition (I<sub>50</sub>) of an Enzymatic Reaction" *Biochem. Pharmacol.* **1973**, *22*, 3099-3108.
- <sup>xvii</sup> Chaires, J. B. "Possible Origin of Differences between van't Hoff and Calorimetric Enthalpy Estimates." *Biophysical Chemistry*, **1997**, *64*, 15-23.
- <sup>xviii</sup> Horn, J. R.; Russell, D.; Lewis, E. A.; Murphy, K. P. "van't Hoff and Calorimetric Enthalpies from Isothermal Titration Calorimetry: Are There Significant Discrepancies?" *Biochemistry* **2001**, *40*, 1774-1778.
- <sup>xix</sup> Horn, J.R.; Brandts, J.F.; Murphy, K.P. "van't Hoff and Calorimetric Enthalpies II: Effects of Linked Equilibria" *Biochemistry*, **2002**, *41*, 7501-7507.
- <sup>xx</sup> Ragone, R.; Colonna, G. "Reliability of the van't Hoff Plots" *J. Phys. Chem.* **1995**, *99*, 13050.
- <sup>xxi</sup> Wiseman, T., S. Williston, J. F. Brandts, and L.-N. Lin. "Rapid measurement of binding constants and heats of binding using a new titration calorimeter." *Anal. Biochem.*, **1989**, *179*, 131–137.
- <sup>xxii</sup> (a) Cooper, A. *Curr. Opin. Chem. Biol.* **1999**, *3*, 557-563. (b) Haq, I.; Ladbury, J.E. *J. Mol. Recognit.* **2000**, *13*, 188-197. (c) Leavitt, S.; Freire, E. *Curr. Opin. Struct. Biol.* **2003**, *11*, 560-566.

- <sup>xxiii</sup> John E. DeLorbe, Dissertation. University of Texas at Austin.
- <sup>xxiv</sup> Indyk, L.; Fisher, H.F. *Methods Enzymol.* **1998**, *295*, 350-364.
- <sup>xxv</sup> Fisher, H. F.; Singh, N. *Methods Enzymol.* **1995**, *259*, 194-221.
- <sup>xxvi</sup> Turnbull, W.B.; Daran, A.H. "On the Value of  $c$ : Can Low Affinity Systems be Studied by Isothermal Titration Calorimetry?" *J. Amer. Chem. Soc.*, **2003**, *125*, 14859-14866.
- <sup>xxvii</sup> Used from ref 26 by permission.
- <sup>xxviii</sup> Willaims, R. J. P. "Flexible Drug Molecules and Dynamic Receptors" *Angew. Chem. Int. Ed. Engl.* **1977**, *16*, 766-777.
- <sup>xxix</sup> Mann, A. In *The Practice of Medicinal Chemistry*, 2 ed; Wermuth, C. G.; Academic Press: London, UK, 2003; pp233-250.
- <sup>xxx</sup> Jenks, W. P. "Binding energy, specificity, and enzymic catalysis: the circe effect." *Adv. Enzymol.* **1975**, *43*, 219-410.
- <sup>xxxi</sup> Fersht, A. *Enzyme Structure and Mechanism*, 2<sup>nd</sup> ed.; Freeman: New York.
- <sup>xxxii</sup> Williams, D.H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, Ute. Kaye, P. T.; Lal, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C. "Toward Semiquantitative Estimation of Binding Constants. Guides for Peptide-Peptide Binding in Aqueous Solution." *J. Am. Chem. Soc.* **1991**, *113*, 7020-7030.
- <sup>xxxiii</sup> Page, M. I.; Jenks, W. P. "Entropic contributions to Rate Accelerations in Enzymatic and Intramolecular Reactions and the Chelate Effect." *Proc. Nat. Acad. Sci. USA*, **1971**, *68*, 1678-1683.
- <sup>xxxiv</sup> O'Neal, H. E.; Benson, S. W. "Entropies and Heat Capacities of Cyclic and Polycyclic Compounds." *J. Chem. Eng. Data* **1970**, *15*, 266-276.
- <sup>xxxv</sup> Pitzer, K. S. "The Vibration Frequencies and Thermodynamic Functions of Long Chain Hydrocarbons." *J. Chem. Phys.*, **1940**, *8*, 711-721.
- <sup>xxxvi</sup> Person, W.B.; Pimental "Thermodynamic Properties and the Characteristic CH<sub>2</sub> Frequencies of  $n$ -Paraffins."

<sup>xxxvii</sup> Gerhard, U.; Searle, M. S.; Williams, D. H. "The Free Energy of Restricting a Bond Rotation in the Binding of Peptide Analogues to Vancomycin Group Antibiotics." *Bioorg. Med. Chem. Lett.* **1993**, *3*, 803-808.

<sup>xxxviii</sup> Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. "Understanding Noncovalent Interactions: Ligand Binding Energy and Catalytic Efficiency from Ligand-Induced Reductions within Receptors and Enzymes" *Angew. Chem. Int. Ed.* **2004**, *43*, 6596-6616.

<sup>xxxix</sup> Andrews, P.R.; Craik, D.J.; Martin, J. L. "Functional Group Contributions to Drug-Receptor Interactions." *J. Med. Chem.* **1984**, *27*, 1648-1657.

<sup>xi</sup> Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. "Molecular Properties That Influence the Oral Bioavailability of Drug Candidates" *J. Med. Chem.* **2002**, *45*, 2615-2623.

<sup>xli</sup> Garcia-Echeverria, C.; Gay, B.; Rahuel, J.; Furet, P. "Mapping the X+1 binding site of the Grb2-SH2 domain with  $\alpha,\alpha$ -disubstituted cyclic  $\alpha$ - amino acids" *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2915-2920.

<sup>xlii</sup> Tyndall, J. D. A. N., T.; Fairlie, D. P. . "Proteases Universally Recognize %-Strands in Their Active Sites" *Chem. Rev.* **2005**, *105*, 973-999.

<sup>xliii</sup> Glenn, M. P.; Fairlie, D. P. "Mimetics of the Peptide  $\beta$  -Strand" *Mini Rev. in Med. Chem.* **2002**, *2*, 433-445.

<sup>xliv</sup> Tyndall, J. D. A. N., T.; Fairlie, D. P. . "Proteases Universally Recognize %-Strands in Their Active Sites" *Chem. Rev.* **2005**, *105*, 973-999.

<sup>xlv</sup> (a) Benfield, A. P.; Teresk, M. G.; Plake, H. R.; DeLorbe, J. E.; Millspaugh, L. E.; Martin, S. F. *Angew. Chem. Int. Edit.* **2006**, *45*, 6830-6835. (b) DeLorbe, J. E.; Clements, J. H.; Teresk, M. G.; Benfield, A. P.; Plake, H. R.; Millspaugh, L. E.; Martin, S. F. *J. Am. Chem. Soc.* **2009**, *131*, 16758-16770 and references therein.

<sup>xlvi</sup> Tronrud, D.E.; Holden, H. M.; Matthews, B. W.; "Structures of Two Thermolysin-Inhibitor Complexes that Differ by a Single Hydrogen Bond" *Science* **1987**, *235*, 571-574.



<sup>xlvi</sup> Morgan, B. P.; Holland, D. R.; Matthews, B. W.; Bartlett, P. A. "Structure-Based Design of an Inhibitor of the Zinc Peptidase Thermolysin." *J. Am. Chem. Soc.* **1994**, *116*, 3251-3260.

<sup>xlvi</sup> Fraser, M.E.; Styrnadka, N.C. J.; Bartlett, P.A.; Hanson, J.E.; James, M.N.G. "Crystallographic analysis of transition-state mimics bound to penicillopepsin: phosphorus-containing peptide analogues." *Biochemistry* **1992**, *31*, 5201.

<sup>xlvi</sup> Smith, W. W.; Bartlett, P. A. "Macrocyclic Inhibitors of Penicillopepsin. 3. Design, Synthesis, and Evaluation of an inhibitor Bridged between P2 and P1'." *J. Am. Chem. Soc.* **2002**, *124*, 4622-4628.

<sup>i</sup> Chen, W.; Chang, C-E.; Gilson, M. "Calculation of Cyclodextrin Binding Affinities: Energy, Entropy and implications for Drug Design." *Biophys. J.* **2004**, *87*, 3035-49.

<sup>ii</sup> Gilson, M. K.; Zhou, H-X. "Calculation of Protein-Ligand Binding Affinities." *Ann. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21-42.

<sup>iii</sup> Freidinger, R. M.; Veber, D. F.; Perow, D. S.; Brooks, J. R.; Saperstein, R. "Bioactive Conformation of Luteinizing Hormone-Releasing Hormone: Evidence from a Conformationally Constrained Analog" *Science* **1980**, *120*, 656-658.

<sup>iii</sup> Eblinger, F.; Schneider, H.-J. "Stabilities of Hydrogen-Bonded Supramolecular Complexes with Various Numbers of Single Bonds: Attempts to Quantify a Dogma in Host-Guest Chemistry." *Angew. Chem. Int. Ed.* **1998**, *37*, 826-829.

<sup>iv</sup> Udugamasooriya, D. G.; Spaller, M. R. "Conformational Constraint in Protein Ligand Design and the Inconsistency of Binding Entropy" *Biopolymers* **2008**, *89*, 653-667.

<sup>iv</sup> Li, T.; Saro, D.; Spaller, M. R. "Thermodynamic Profiling of Conformationally Constrained Cyclic Ligands for the PDZ Domain" *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1385-1388.

<sup>vi</sup> Harris, B. Z.; Lim, W. A. "Mechanism and Role of PDZ Domains in Signaling Complex Assembly." *J. Cell. Sci.* **2001**, *114*, 3219-3231.

<sup>vii</sup> Sheng, M.; Sala, "PDZ Domains and the Organization of Supramolecular Complexes." *C. Ann. Rev. Neurosci.* **2001**, *24*, 1-29.

- <sup>lviii</sup> Doyle, D. A.; Lee, A.; Lewis, J.; Kim, E.; Sheng, M.; MacKinnon, R. "Crystal Structures of a Complexed and Peptide-Free Membrane Protein-Binding Domain: Molecular Basis of Peptide Recognition by PDZ" *Cell* **1996**, 85, 1067–1076.
- <sup>lix</sup> Udugamasooriya, D. G.; Spaller, M. R. "Conformational Constraint in Protein Ligand Design and the Inconsistency of Binding Entropy" *Biopolymers* **2008**, 89, 653-667.
- <sup>lx</sup> Widlanski, T.; Bender, S. L.; Knowles, J. R. *J. Am. Chem. Soc.* **1989**, 111, 2299-2300.
- <sup>lxi</sup> Harrison, B. A.; Gierasch, T. M.; Neilan, C.; Pasternak, G. W.; Verdine, G. L. "High-Affinity Mu Opioid Receptor Ligands Discovered by the Screening of an Exhaustively Stereodiversified Library of 1,5-Enediols." *J. Am. Chem. Soc.* **2002**, 124, 13352-13353.
- <sup>lxii</sup> Martin, S.F.; Austin, R.E.; Oalman, C.J. "Stereoselective Synthesis of 1,2,3-Trisubstituted Cyclopropanes as Novel Dipeptide Isosteres." *Tetrahedron Lett.* **1990**, 31, 4731-4734.
- <sup>lxiii</sup> Reichelt, A. and Martin, S. F. "Synthesis and Properties of Cyclopropane-Derived Peptidomimetics." *Acc. Chem. Res.* **2006**, 39, 433-442.
- <sup>lxiv</sup> Martin, S. F.; Dwyer, M. P.; Hartmann, B.; Knight, K. S. "Cyclopropane-Derived Peptidomimetics. Design, Synthesis and Evaluation of Novel Enkephalin Analogues." *J. Org. Chem.* **2000**, 65, 1305-1318.
- <sup>lxv</sup> Martin, S.F.; Austin, R. E.; Oalman, C. J.; Baker, W.R.; Condon, S. L.; deLara, E.; Rosenberg, S.H.; Spina, K.P.; Stein, H. H.; Cohen, J.; Kleinert, H. D. "1,2,3-Trisubstituted Cyclopropanes as Conformationally Restricted Peptide Isosteres: Application to the Design and Synthesis of Novel Renin Inhibitors" *J. Med. Chem.* **1992**, 35, 1710-1721.
- <sup>lxvi</sup> Martin, S.F.; Oalman, C. J.; Liras, S. "Cyclopropanes as Conformationally restricted peptide isosteres. Design and Synthesis of Novel Collagenase Inhibitors" *Tetrahedron*, **1993**, 49, 3521–3532.
- <sup>lxvii</sup> Reichelt, A.; Gaul, C.; Frey, R.R.; Kennedy, A.; Martin, S.F. "Design, Synthesis, and Evaluation of Matrix Metalloprotease Inhibitors Bearing Cyclopropane-Derived Peptidomimetics as P1' and P2' Replacements. *J. Org. Chem.* **2002**, 67, 4062-4075.

<sup>lxxviii</sup> For references on enthalpy/entropy compensation, see: (a) Krishnamurthy, V.; Bohall, B.; Semetey, V.; Whitesides, G. *J. Am. Chem. Soc.* "The Paradoxical Thermodynamic Basis for the Interaction of Ethylene Glycol, Glycine, and Sarcosine Chains with Bovine Carbonic Anhydrase II: An Unexpected Manifestation of Enthalpy/Entropy Compensation" **2006**, *128*, 5802. (b) Krug, R. R.; Hunter, W. G.; Grieger, R. A. *The Journal of Physical Chemistry* "Enthalpy-entropy compensation. 1. Some fundamental statistical problems associated with the analysis of van't Hoff and Arrhenius data" **1976**, *80*, 2335. (c) Olsson, T. S. G.; Ladbury, J. E.; Pitt, W. R.; Williams, M. A. *Protein Sci.* "Extent of enthalpy-entropy compensation in protein-ligand interactions" **2011**, *20*, 1607.

<sup>lxxix</sup> Davidson, J.P. Ph.D. Dissertation, The University of Texas, 2001.

<sup>lxxx</sup> Davidson, J.P.; Lubman, O.; Rose, T.; Waksman, G. Martin, S. F. "Calorimetric and Structural Studies of 1,2,3-Trisubstituted Cyclopropanes as Conformationally Constrained Peptide Inhibitors of Src SH2 Domain Binding" *J. Amer. Chem. Soc.* **2002**, *124*, 205–215.

<sup>lxxxi</sup> Karplus, M.; McCammon, J. A. "Molecular dynamics simulations of biomolecules." *Nat. Struct. Biol.* **2002**, *9*, 646–652.

<sup>lxxxii</sup> Karplus, M.; Ichiye, T.; Pettit, B. M. "Configurational entropy of native proteins." *Biophys. J.* **1987**, *52*, 1083–1085.

<sup>lxxxiii</sup> Waksman, G.; Shoelson, S. E.; Pant, N.; Cowburn, D.; Kuriyan, J. "Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms." *Cell* **1993**, *72*, 779–790.

<sup>lxxxiv</sup> Ward, J.M.; Gorenstein, N.M.; Tian, J.; Martin, S.F.; Post, C.B. "Constraining binding hot spots: NMR and molecular dynamics simulations provide a structural explanation for enthalpy-entropy compensation in SH2-ligand binding." *J. Am. Chem. Soc.* **2010**, *132*, 11058–11070.

<sup>lxxxv</sup> Benfield, A.M; Teresk, M.G.; Plake, H.R.; Delorbe, J.E.; Millspaugh, L.E.; Martin, S.F. "Ligand Preorganization May be Accompanied by Entropic Penalties in Protein-Ligand Interactions." *Angew. Chem. Int. Ed.* **2006**, *45*, 6830–6834.

<sup>lxxxvi</sup> DeLorbe, J.; Clements, J.; M. Teresk; A. Benfield; H. Plake, L.; Millspaugh, L. and Martin, S.F. "Thermodynamic and Structural Effects of Conformational Constraints in Protein-Ligand Interactions. Entropic Paradoxy Associated with Ligand Preorganization." *J. Amer. Chem. Soc.* **2009**, *131*, 16758.

- <sup>lxxvii</sup> Shi, Y.; Zhu, C. Z.; Martin, S.F.; Ren, R. "Probing the Effect of Conformational Constraint on Phosphorylated Ligand Binding to an SH2 Domain Using Polarizable Force Field Simulations" *J. Phys. Chem. B.* **2012**, *116*, 1716-1727.
- <sup>lxxviii</sup> Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery" *Advanced Drug Delivery reviews* **1997**, *23*, 3-25.
- <sup>lxxix</sup> Ruben, A. J.; Kiso, Y.; Freire, E. "Overcoming Roadblocks in Lead Optimization: A Thermodynamic Prospective" *Chem. Biol. Drug. Des.* **2006**, *67*, 2-4.
- <sup>lxxx</sup> Garcia-Echeverria, C.; Gay, B.; Rahuel, J.; Furet, P. "Mapping the X+1 binding site of the Grb2-SH2 domain with  $\alpha,\alpha$ -disubstituted cyclic  $\alpha$ - amino acids" *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2915-2920.
- <sup>lxxxi</sup> Furet, P.; Gay, B.; Caravatti, G.; Garcia-Echeverria, C.; Fretz, H.; Rahuel, J.; "Structure-based design of peptidomimetic ligands of the Grb2-SH2 domain" *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2865-2870.
- <sup>lxxxii</sup> Garica-Echevaria, C.; Gay, B; Rahuel, J.; Furet, P. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2915-2920.
- <sup>lxxxiii</sup> Myslinski, J.M.; Delorbe, J. E.; Clements, J.C.; Martin, S.F. "Protein-Ligand Interactions: Thermodynamic Effects Associated with Increasing Nonpolar Surface Area" *J. Amer. Chem. Soc.* **2011**, *133*, 18518-18521.
- <sup>lxxxiv</sup> James M. Myslinski, dissertation. University of Texas at Austin.
- <sup>lxxxv</sup> Schoepfer, J.; Fretz, H. Gay, B.; Furet, P.; Garcia-Echevaria, C.; End, N.; Caravatti, G. "Highly Potent Inhibitors of the Grb2-SH2 Domain" *Bioorg. Med. Chem. Lett.* **1999**, *9*, 221-226.
- <sup>lxxxvi</sup> Furet, P.; Gay, B.; Caravatti, G.; Garcia-Echevaria, C.; Rahuel, J.; Schoepfer, J.; Fretz, H. "Structure-Based Design and Synthesis of High Affinity Tripeptide Ligands of the Grb2-SH2 Domain" *J. Med. Chem.* **1998**, *41*, 3442-3449
- <sup>lxxxvii</sup> Olsson, T. S. G.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. "The Thermodynamics of Protein-Ligand Interaction and Solvation: Insights for Ligand Design" *J. Mol. Biol.* **2008**, *384*, 1002-1017.

Faergemna, N. J.; Sigurskjold, B. W.; Kragelund, B. B.; Andersen, K. V.; Knudsen, J. "Thermodynamics of Ligand Binding to Acyl-Coenzyme A Binding Protein Studied by Titration Calorimetry" *Biochemistry*, **1996**, 35, 14118–14126.

<sup>lxxxviii</sup> For reviews, see: a.) Schneider, H.J. "Binding mechanisms in supramolecular complexes" *Angew. Chem. Int. Ed. Engl.*, **2009**, 48, 3924-3977. b.) Blokzijl, W.; Engberts, J. B. F. N. "Hydrophobic effects. Opinion and facts" *Angew. Chem. Int. Ed. Engl.*, **1993**, 11, 1545-1579.

<sup>lxxxix</sup> Ferguson, S.B.; Seward, E.M.; Diederich, F.; Sanford, E.M.; Chou, A.; Inocencio-Szweda, P.; Knobler, C.B.; "Strong enthalpically driven complexation of neutral benzene guests in aqueous solution" *J. Org. Chem.* **1988**, 53, 5593-5595.

<sup>xc</sup> Ross, P.; Subramanian, S. "Thermodynamics of protein association reactions: forces contributing to stability" *Biochemistry*, **1981**, 20, 3096-3102.

<sup>xc<sup>i</sup></sup> Banerjee, S.K.; Pogolotti, A.; Rupley, J.; "Self-association of lysozyme. Thermochemical measurements: effect of chemical modification of Trp-62, Trp-108, and Glu-35." *J. Biol. Chem.*, **1975**, 250, 8260-8266.

<sup>xc<sup>ii</sup></sup> Velick, S.F.; Baggott, J.P.; Sturtevant, J.M. "Thermodynamics of nicotinamide-adenine dinucleotide addition to the glyceraldehyde 3-phosphate dehydrogenases of yeast and of rabbit skeletal muscle. An equilibrium and calorimetric analysis over a range of temperatures." *Biochemistry*, **1971**, 10, 779–786.

<sup>xc<sup>iii</sup></sup> Gerlach, C.; Smolinski, M.; Steuber, H.; Sotriffer, C. A.; Heine, A. Hangauer, D. C.; Klebe, G. "Thermodynamic Inhibition Profile of a Cyclopentyl and a Cyclohexyl Derivative towards Thrombin: The Same but for Different Reason" *Angew. Chem., Int. Ed.* **2007**, 46, 8511-8514.

<sup>xc<sup>iv</sup></sup> Englert L, Biela A, Zayed M, Heine A, Hangauer D, Klebe G. 2010. Displacement of disordered water molecules from hydrophobic pocket creates enthalpic signature: binding of phosphoramidate to the S1'-pocket of thermolysin. *Biochim Biophys Acta*. 1800:1192-1202.

<sup>xc<sup>v</sup></sup> Matthews, B. W.; Colman, P.M.; Jansonius, J.N.; Titani, K.; Walsh, K. A.; Neurath, W. "Structure of Thermolysin" *Nat. New Biol.* **1972**, 238, 41-43.

- <sup>xcvi</sup> Patchett, A. A.; Cordes, E. H. "The Design and Properties of N-carboxylalkyldipeptide Inhibitors of Angiotensin-Converting Enzyme" *Adv. Enzymol. Relat. Areas Mol. Biol.* **1985**, 57, 1-84.
- <sup>xcvii</sup> Young, T. Y.; Hua, L.; Huang, X.; Abel, R.; Friesner, R.; Berne, B. J. "Dewetting Transitions in Protein Cavities" *Proteins*, **2010**, 78, 1856-1869.
- <sup>xcviii</sup> Bingham, R.; Findlay, J. B.; Hsieh, S-y; Kalverda, A. P.; Kjielberg, A.; Perazzolo, C.; Phillips, S. E.; Seshadri, K.; Trinh, C. H. Turnbull, W. B.; Bodenhausen, G.; Homans, S. W. "Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein." *J. Am. Chem. Soc.* **2004**, 126, 1675-1681.
- <sup>xcix</sup> Sharrow, S. D.; Novotny, M. V.; Stone, M. J. "Thermodynamics of Binding between Mouse Major Urinary Protein-I and the Pheromone 2-sec-Butyl-4,5-dihydrothiazole" *Biochemistry*, **2003**, 42, 6302-6309.
- <sup>c</sup> Talhout, R.; Villa, A.; Mark, A. E.; Engberts, J. B. F. N. Understanding Binding Affinity: A Combined Isothermal Titration Calorimetry/Molecular Dynamics Study of the Binding of a Series of Hydrophobically Modified Benzamidinium Chloride Inhibitors to Trypsin" *J. Am. Chem. Soc.* **2003**, 125, 10570-10579.
- <sup>ci</sup> Carey, C.; Cheng, Y.-K.; Rossky, P. "Hydration structure of the chymotrypsin substrate binding pocket: the impact of constrained geometry" *Chem. Phys.* **2000**, 258, 415-425.
- <sup>cii</sup> Used from reference ci with permission.
- <sup>ciii</sup> Kawasaki, Y.; Freire, E. "How Much Binding Affinity can be Gained by Filling a Cavity?" *Chem. Biol. Drug Des.* **2010**, 75, 143-151.
- <sup>civ</sup> Kawasaki, Y.; Freire, E. "Finding a Better Path to Drug Selectivity." *Drug Discovery Today*, **2011**, 16, 985-999.
- <sup>cv</sup> Memic, A.; Spaller, M. R. "How Do Halogen Substituents Contribute to Protein-Binding Interactions? A Thermodynamic Study of Peptide Ligands with Diverse Aryl Halides" *ChemBioChem* **2008**, 2793-2795.
- <sup>cvi</sup> Bondi, A. "van der Waals Volumes and Radii." *J. Phys. Chem.* **1964**, 68, 441-451.

<sup>cvi</sup> Mecinovic, J.; Snyder, P.A.; Mirica, K. A.; Bai, S.; Mack, E.T.; Kwant, R. L.; Moustakas, D. T.; Heroux, A.; Whitesides, G.M. "Fluoroalkyl and alkyl chains have similar hydrophobicities in binding to the "hydrophobic wall" of carbonic anhydrase." *J. Am. Chem. Soc.* 2011, *133*, 14017-14026.

<sup>cvi</sup> Topliss, J.G. "Utilization of Operational Schemes for Analog Synthesis in Drug Design." *J. Med. Chem.* **1972**, *10*, 1006-1012.

<sup>cix</sup> Fujita, T.; Iwasa, J.; Hansh, C. "A New Substituent Constant,  $\pi$ , Derived from Partition Coefficients." *J. Amer. Chem. Soc.* **1964**, *86*, 5173-5174.

<sup>cx</sup> Hansch, C.; Leo, A.; Unger, S. H-Kim, K.; Nikaitani, D.; Lien, E. J. "'Aromatic" Substituent Constants for Structure-Activity Correlations." *J. Med. Chem.* **1972**, *16*, 1207-1216.

<sup>cx</sup> Obtained from <sup>cvi</sup>, used by permission.

<sup>cxii</sup> For some examples, see: (a) Wang, Z.; Shi, X.H.; Wang, J.; Zhou, T.; Xu, Y. Z.;Huang, T.T.; Li, Y-F.; Zhao, Y.-L.; Yang, L.; Yang, S-Y.; Yu, L-T.; Wei, Q-Y. "Synthesis, Structure-activity relationships and preliminary antitumor evaluation of benzothiazole-2-thiol derivatives as novel apoptosis inducers." *Bioorg. Med. Chem. Lett.*, **2011**, *21*, 1097-1101. (b) Keating, T.A. "Discovery of Selective and Potent Inhibitors of Gram-Positive Bacterial Thymidylate Kinase (TMK)." *J. Med. Chem.* **2012**, *55*,10010–10021. (c) Mills, J. E. *Med. Chem. Commun.*, **2012**, *3*, 174-178.

<sup>cxiii</sup> Wone, D. W. G.; Rowley, J. A.; Garofalo, A. W.; Berkman, C. E. "Optimizing Phenylethylphosponamidates for the Inhibition of Prostate-Specific Membrane Antigen." **2006**, *14*, 67-76.

<sup>cxiv</sup> Benfield, A.M; Teresk, M.G.; Plake, H.R.; Delorbe, J.E.; Millspaugh, L.E.; Martin, S.F. "Ligand Preorganization May be Accompanied by Entropic Penalties in Protein-Ligand Interactions." *Angew. Chem. Int. Ed.* **2006**, *45*, 6830-6834.

<sup>cxv</sup> DeLorbe, J.; Clements, J.; M. Teresk; A. Benfield; H. Plake, L.; Millspaugh, L. and Martin, S.F. "Thermodynamic and Structural Effects of Conformational Constraints in Protein-Ligand Interactions. Entropic Paradoxy Associated with Ligand Preorganization." *J. Amer. Chem. Soc.* **2009**, *131*, 16758.

- <sup>cxvi</sup> Shi, Y.; Zhu, C. Z.; Martin, S.F.; Ren, R. "Probing the Effect of Conformational Constraint on Phosphorylated Ligand Binding to an SH2 Domain Using Polarizable Force Field Simulations" *J. Phys. Chem. B.* **2012**, *116*, 1716-1727.
- <sup>cxvii</sup> Flower, D.R. "The lipocalin protein and family: Structure and function." *Biochem. J.* **1996**, *318*, 1-14.
- <sup>cxviii</sup> Bingham, R.J. *J. Am. Chem. Soc.*; **2004**, *126*, 1675-1681.
- <sup>cxix</sup> Pertinhez, T. A.; Ferrari, E.; Casali, E.; Patel, J. A.; Spinsi A.; Smith, L. J. *Biochemical and Biophysical research Communications*, **2009**, *390*, 1266-1271.
- <sup>cxx</sup> Wang, L.; Berne, B. J.; Friesner, R. A. "Ligand binding to protein-binding pockets with wet and dry regions" *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1326-1330.
- <sup>cxxi</sup> According to literature procedure.
- <sup>cxxii</sup> Martin, S.F.; Austin, R.E.; Oalman, C.J.; Baker, W.R.; Condon, E., Delara, S.H.; Rosenberg, K.P.; Spina, H. H.; Stein, J. C.; Kleinert, H. D. "1,2,3-Trisubstituted Cyclopropanes as Conformationally Restricted Peptide Iossteres: Application to the Design and Synthesis of Novel Renin Inhibitors." *J. Med. Chem.* **1992**, *35*, 1710-1721.
- <sup>cxxiii</sup> Davidson, J. P.; Lubman, O.; Rose, T.; Waksman, G.; Martin, S. F. "Calorimetric and structural studies of 1,2,3-trisubstituted cyclopropanes as conformationally constrained peptide inhibitors of Src SH2 domain binding" *J. Am. Chem. Soc.* **2002**, *124*, 205.
- <sup>cxxiv</sup> Corey, E. J.; Myers, A. G. "Efficient Synthesis and Intramolecular Cyclopropanation of Unsaturated Diazoacetic Esters" *Tetrahedron Lett.* **1984**, *25*, 3559-3562.
- <sup>cxxv</sup> Martin, S.F.; Austin, R. E.; Oalman, C. J.; Baker, W. R.; Condon, S. L.; Delara, E., Rosenberg, S. H.; Spina, K. P.; Stein, H. H. "1,2,3-Trisubstituted Cyclopropanes as Conformationally Restricted Isosteres: Application to the Design and Synthesis of Novel Renin Inhibitors" *J. Med. Chem.*, **1992**, *35*, 1710-1721.



<sup>cxxvi</sup> Martin, S.F.; Dwyer, M. P.; Lynch, C.L. "Application of AlMe<sub>3</sub>-Mediated Amidation Reactions for Peptide Synthesis." *Tetrahedron Lett.* **1998**, 39, 1517-1520.

<sup>cxxvii</sup> Malham, R.; Johnstone, S.; Bingham, R.J.; Barratt, E.; Phillips, S. E. V.; Laughton, C. A.; Homans, S. W. "Strong Solute-Solute Dispersive Interactions in a Protein-Ligand Complex" *J. Am. Chem. Soc.* **2005**, 127, 640-647.

<sup>cxxviii</sup> . (a) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. USA* **1971**, 68, 1678-1683. (b) Gillet, V.; Johnson, A. P.; Mata, P.; Sike, S.; Williams, P. *J. Comput. Aided. Mol. Des.* **1993**, 7, 127-153. (c) Raha, K.; Merz, K. M. *J. Med. Chem.* **2005**, 48, 4558-4575. (d) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Biophys. Biomol.* **2007**, 36, 21-42. (e) Gao, C.; Park, M. S.; Stern, H. A. *Biophys. J.* **2010**, 98, 901-910.

<sup>cxxix</sup> Gilson, M. K.; Zhou, H-X. "Calculation of Protein-Ligand Binding Affinities." *Ann. Rev. Biophys. Biomol. Struct.* **2007**, 36, 21-42.

<sup>cxxx</sup> Chen, W.; Chang, C-E.; Gilson, M. "Calculation of Cyclodextrin Binding Affinities: Energy, Entropy and implications for Drug Design." *Biophys. J.* **2004**, 87, 3035-49.

<sup>cxxxi</sup> Charette, A. B.; Molinaro, C.; Brochu, C. "Catalytic Asymmetric Cyclopropanation of Allylic Alcohols with Titanium-TADDOLate: Scope of the Cyclopropanation Reaction" *J. Am. Chem. Soc.* **2001**, 123, 12168-12175.

<sup>cxxxii</sup> Kim, I.S.; Dong, G. R.; Hoon, Y. J; "Palladium(II)-Catalyzed Isomerization of Olefins with tributyltin Hydride" *J. Org. Chem.* **2007**, 72, 5424-5426.

<sup>cxxxiii</sup> Grünanger, C. U.; Breit, B. "Branched-Regioselective Hydroformylation with Catalytic Amounts of a Reversibly Bound Directing Group" *Angew. Chem. Int. Ed.* **2008**, 47, 7346-7349.

<sup>cxxxiv</sup> : Spivey, A. C.; Laraia, L.; Bayly, A. R.; Rzepa, H. S.; White, A. J. P. "Stereoselective Synthesis of *cis*- and *trans*-2,3-Disubstituted Tetrahydrofuran via Oxonium-Prins Cyclization: Access to the Cordigol Ring System" *Org. Lett.* **2010**, 12, 900-903.

<sup>cxv</sup> Hunter, C. A. “Quantifying Intermolecular Interactions: Guidelines for the Molecular Recognition Toolbox” *Angew. Chem. Int. Ed.* **2004**, *43*, 5310-5324.

<sup>cxvi</sup> Talhout, R.; Villa, A.; Mark, A. E.; Enberts, J. B. F. N. *J. Am. Chem. Soc.* **2003**, *127*, 10570-10579.

<sup>cxvii</sup> Mecinovic, J.; Snyder, P.A.; Mirica, K. A.; Bai, S.; Mack, E.T.; Kwant, R. L.; Moustakas, D. T.; Heroux, A.; Whitesides, G.M. *J. Am. Chem. Soc.* **2011**, *133*, 14017-14026.

<sup>cxviii</sup> Widlanski, T.; Bender, S. L.; Knowles, J. R. “Dehydroquinase Synthase: a Sheep in Wolf’s Clothing” *J. Am. Chem. Soc.* **1989**, *111*, 2299-2300.

<sup>cxix</sup> Gerhard, U.; Searle, M. S.; Williams, D. H. “The Free Energy of Restricting a Bond Rotation in the Binding of Peptide Analogues to Vancomycin Group Antibiotics.” *Bioorg. Med. Chem. Lett.* **1993**, *3*, 803-808.

<sup>cxl</sup> Harrison, B. A.; Gierasch, T. M.; Neilan, C.; Pasternak, G. W.; Verdine, G. L. *J. Am. Chem. Soc.* **2002**, *124*, 13352-13353.

<sup>cxli</sup> Dunitz, J. D. *Science* **1994**, *264*, 670.

<sup>cxlii</sup> Homans, S. W. *Drug Discovery Today* **2007**, *12*, 534-539.

<sup>cxliii</sup> (a) Cheng, Y.-K.; Rossky, P. J. *Nature* **1998**, *392*, 696-699. (b) Carey, C.; Cheng, Y.-K.; Rossky, P. *Chem. Phys.* **2000**, *258*, 415-425. (c) Li, Z.; Lazaridis, T. *J. Phys. Chem. B* **2005**, *109*, 662-670.

<sup>cxliv</sup> Plyasunov, A. V.; Shock, E. L. *Geochim. Cosmochim. Acta.* **2000**, *64*, 439-468

<sup>cxlv</sup> (a) Böcskei, Z.; Groom, C. R.; Flower, D. R.; Wright, C. E.; Philips, S. E. V.; Cavaggioni, A.; Findlay, J. B. C.; North, A. C. T. *Nature* **1992**, *360*, 186-188. (b) Timm, D. E.; Baker, L. J.; Mueller, H.; Zidek, L.; Novotny, M. V. *Protein Sci.* **2001**, *10*, 997-1004.

<sup>cxlvi</sup> Wang, L.; Berne, B. J.; Friesner, R. A. “Ligand binding to protein-binding pockets with wet and dry regions” *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1326-1330.

<sup>cxlvii</sup> Tuñón, I.; Silla, E.; Pascual-Ahjuir, J. L. *Protein Eng.* **1992**, *5*, 715-716.

- <sup>cxlviii</sup> Yang, C.; Wang, R.; Wang, S. "A Systematic Analysis of the Effects of Small-Molecule Binding on Protein Flexibility of the Ligand-Binding Sites" *J. Med. Chem.* **2005**, *48*, 5648-5650.
- <sup>cxlix</sup> Ringe, D.; Petsko, G. A. *Methods Enzymol.* **1986**, *131*, 389-433.
- <sup>cl</sup> Tuñón, I.; Silla, E.; Pascual-Ahjuir, J. L. *Protein Eng.* **1992**, *5*, 715-716.
- <sup>cli</sup> Myslinski, J. M.; DeLorbe, J. E.; Clements, J. H.; Martin, S. F. *J. Amer. Chem. Soc.* **2011**, *133*, 18518-18521.
- <sup>clii</sup> Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery" *Advanced Drug Delivery reviews* **1997**, *23*, 3-25.
- <sup>cliii</sup> Ruben, A. J.; Kiso, Y.; Freire, E. "Overcoming Roadblocks in Lead Optimization: A Thermodynamic Prospective" *Chem. Biol. Drug Des.* **2006**, *67*, 2-4.
- <sup>cliv</sup> Kawasaki, Y.; Freire, E. "How Much Binding Affinity can be Gained by Filling a Cavity?" *Chem. Biol. Drug Des.* **2010**, *75*, 143-151.
- <sup>clv</sup> Myslinski, J.M.; Delorbe, J. E.; Clements, J.C.; Martin, S.F. "Protein-Ligand Interactions: Thermodynamic Effects Associated with Increasing Nonpolar Surface Area" *J. Amer. Chem. Soc.* **2011**, *133*, 18518-18521.
- <sup>clvi</sup> Fretz, H.; Furet, P.; Garcia-Echeverria, C.; Rahuel, J.; Schoepfer, J. "Structure-based design of compounds inhibiting Grb2-SH2 mediated protein-protein interactions in signal transduction pathways" *Curr. Pharm. Des.* **2000**, *6*, 1777-1796.
- <sup>clvii</sup> Furet, P.; Gay, B.; Caravatti, G; García-Echeverría, C; Rahuel, J; Schoepfer, J and Fretz, H. "Structure-based design and synthesis of high affinity tripeptide ligands of the Grb2-SH2 domain" *J. Med. Chem.* **1998**, *41*, 3442.
- <sup>clviii</sup> Cantley, L. C.; Auger, K. R.; Carpenter, C.; Duckworth, B.; Graziani, A.; Kapeller, R.; Soltoff, S. "Oncogenes and Signal Transduction" *Cell*, **1991**, *64*, 281-303.
- <sup>clix</sup> Lowenstein, E. J.; Daly, R. J.; Batzer, A. G.; Li, W.; Margolis, B.; Lammers, R.; Ullrich, A.; Sklonik, E. Y.; Sagi, D. B.; Schlessinger, J. "The SH2 and SH3

Domain-Containing Protein GRB2 Links Receptor Tyrosine Kinases to ras Signaling". *Cell*, **1992**, 431-442.

<sup>clx</sup> Gishizky, M.L.; *Ann. Rep. Med. Chem. Chapt.* **1995**, 26, 247-253.

<sup>clxi</sup> Birchmeier, C.; Birchmeier, W.; Gherardi, E.; Vande Woude V. F. "Met, Metastasis, Motility and More." *Nature Reviews Molecular Cell Biology*, **2003**, 4, 915-925.

<sup>clxii</sup> Gishizky, M.L.; *Ann. Rep. Med. Chem. Chapt.* **1995**, 26, 247-253.

<sup>clxiii</sup> Garcia-Echevaria, *Bioorg. Med. Chem. Lett.*, **1999**, 9, 2915-2920.

<sup>clxiv</sup> Plyasunov, A. V.; Shock, E. L. *Geochim. Cosmochim. Acta.* **2000**, 64, 439-468.

<sup>clxv</sup> Chervenak, M. C., Toone, E. J. "A direct measure of the contribution of solvent reorganization to the enthalpy of ligand binding" *J. Amer. Chem. Soc.* **1994**, 116, 10533-10539.

<sup>clxvi</sup> Topliss, J.G. "*J. Med. Chem.* **1972**, 10, 1006-1012.

<sup>clxvii</sup> Fujita, T.; Iwasa, J.; Hansh, C. "A New Substituent Constant,  $\pi$ , Derived from Partition Coefficients." *J. Amer. Chem. Soc.* **1964**, 86, 5173-5174.

<sup>clxviii</sup> Hansch, C.; Leo, A.; Unger, S. H-Kim, K.; Nikaitani, D.; Lien, E. J. "'Aromatic" Substituent Constants for Structure-Activity Correlations." *J. Med. Chem.* **1972**, 16, 1207-1216.

<sup>clxix</sup> Memic, A.; Spaller, M. R. "How Do Halogen Substituents Contribute to Protein-Binding Interactions? A Thermodynamic Study of Peptide Ligands with Diverse Aryl Halides" *ChemBioChem* **2008**, 2793-2795.

<sup>clxx</sup> Jenks, W.P. "On the Attribution and Additivity of Binding Energetics" *Proc. Nat. Acad. Sci. U.S.A* **1981**, 78, 4046-4050.

<sup>clxxi</sup> James M. Myslinski, Ph.D. dissertation. University of Texas at Austin.

<sup>clxxii</sup> Gao, C.; Park, M. S.; Stern, H. A. "Accounting for Ligand Conformational Restriction in Calculations of Protein-Ligand Binding Affinities" *Biophys. J.* **2010**, 98, 901-910

<sup>clxxiii</sup> Gilson, M. K.; Zhou, H. X. "Calculation of protein-ligand binding affinities." *Annu. Rev. Bioph. Biom.* **2007**, 36, 21-42.

<sup>clxxiv</sup> Raha, K.; Merz, K. M. "Large-Scale Validation of a Quantum Mechanics Based Scoring Function: Predicting the Binding Affinity and the Binding Mode of a Diverse Set of Protein–Ligand Complexes." *J. Med. Chem.* **2005**, *48*, 4558-4575.

<sup>clxxv</sup> Wang, L.; Berne, B. J.; Friesner, R. A. "On achieving high accuracy and reliability in the calculation of relative protein–ligand binding affinities." *Proc Natl Acad Sci USA*, **2012**, *109*, 1937-1942.

<sup>clxxvi</sup> For reports on limitations of scoring functions, see: : (a) Warren, G.L.; Andrews, C. W.; Capello, A.M.; Clarke, B.; LaLonde, J.; Lambert, M.H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Semger, S.; Tedesco, G.; Wall, I.D.; Woolven, J.M.; Peishoff, C.E.; Head, M.S. *J. Med. Chem.* **2006**, *11*, 580-594. (b) Klebe, G.; *Drug Discovery Today* **2006**, *11*, 580-594. (c) Damn, K.L.; Carlson, H. A. *J. Am. Chem. Soc.* **2007**, *129*, 8225-8235.

<sup>clxxvii</sup> Topliss, J. G. "A Manual Method for Applying the Hansch Approach to Drug Design." *J. Med. Chem.* **1977**, *20*, 463-469.

<sup>clxxviii</sup> Topliss, J.G. "Utilization of Operational Schemes for Analog Synthesis in Drug Design." *J. Med. Chem.* **1972**, *10*, 1006-1012.

<sup>clxxix</sup> (a) Keating, T.A. "Discovery of Selective and Potent Inhibitors of Gram-Positive Bacterial Thymidylate Kinase (TMK)" *J. Med. Chem.*, **2012**, *55*, 10010–10021. (b) Mills, J. E. "SAR mining and its application to the design of TRPA1 Antagonists" *Med. Chem. Commun.*, **2012**, *3*, 174-178.

<sup>clxxx</sup> Steinbaugh, B.A.; Hamilton, H.W.; Prasad, V.; Para, K.S.; Tummino, P.J.; Ferguson, D.; Lunney, E.A.; Blankley C.J. "A Topliss Tree Analysis of the HIV-protease Inhibitory Activity of 6-phenyl-4-hydroxy-pyran-2-ones." *Biorg. Med. Chem. Lett.*, **1996**, *10*, 1099-1104.

<sup>clxxxi</sup> Manuscript submitted

<sup>clxxxii</sup> (a) Krishnamurthy, V.; Bohall, B.; Semetey, V.; Whitesides, G. *J. Am. Chem. Soc.* "The Paradoxical Thermodynamic Basis for the Interaction of Ethylene Glycol, Glycine, and Sarcosine Chains with Bovine Carbonic Anhydrase II: An

Unexpected Manifestation of Enthalpy/Entropy Compensation" **2006**, 128, 5802. (b) Krug, R. R.; Hunter, W. G.; Grieger, R. A. *The Journal of Physical Chemistry* "Enthalpy-entropy compensation. 1. Some fundamental statistical problems associated with the analysis of van't Hoff and Arrhenius data" **1976**, 80, 2335. (c) Olsson, T. S. G.; Ladbury, J. E.; Pitt, W. R.; Williams, M. A. *Protein Sci.* "Extent of enthalpy–entropy compensation in protein–ligand interactions" **2011**, 20, 1607.

clxxxiii Shimokhina, N.; Bronowska, A.; Homans, P. S. W. *Angew. Chem. Int. Edit.* **2006**, 45, 6374-6376.

clxxxiv Fujita, T.; Iwasa, J.; Hansch, C. "A New Substituent Constant,  $\pi$ , Derived from Partition Coefficients." *J. Amer. Chem. Soc.* **1964**, 86, 5173-5174.

clxxxv Malham, R.; Johnstone, S.; Bingham, R. J.; Barratt, E.; Phillips, S. E.; Laughton, C. A.; Homans, S. W. *J. Am. Chem. Soc.* **2005**, 127, 17061-17067.

clxxxvi Sharrow, S. D.; Novotny, M. V.; Stone, M. J. "Thermodynamics of Binding between Mouse Major Urinary Protein-I and the Pheromone 2-sec-Butyl-4,5-dihydrothiazole" *Biochemistry*, **2003**, 42, 6302-6309.

clxxxviii Ichihara, O.; Barker, J.; Law, R. J.; Whittaker, M. "Compound Design by Fragment-Linking." *Mol. Inf.* **2011**, 30, 298-306.

clxxxix Jenks, W. P. "On the Attribution and Additivity of Binding Energies" *Proc. Natl. Acad. Sci. USA*, **1981**, 78, 4046-4050.

cx<sup>c</sup> Baum, B.; Muley, L.; Smolinski, M.; Heine, A.; Hangauer, D.; Klebe, G. "Non-additivity of Functional Contributions in Protein-Ligand Binding: A Comprehensive Study by Crystallography and Isothermal Titration Calorimetry." *J. Mol. Biol.* **2010**, 397, 1042-1054.

cxci Tuñón, I.; Silla, E.; Pascual-Ahuir, J. L. "Molecular surface area and hydrophobic effect." *Protein Eng.* **1992**, 5, 715-716.

cxcii Memic, A.; Spaller, M. R. "How Do Halogen Substituents Contribute to Protein-Binding Interactions? A Thermodynamic Study of Peptide Ligands with Diverse Aryl Halides" *ChemBioChem* **2008**, 2793-2795.

<sup>exciii</sup> Bondi, A. "van der Waals Volume and Radii." *J. Phys. Chem.* **1964**, 68, 441-451.

<sup>exciv</sup> Krishnamurthy, V. M.; Bohall, B.; Kim, C. Y.; Moustakas, D.; Christianson, D.; Whitesides, G. M. "Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein-Ligand Binding" *Chem. Rev.* **2008**, 108, 946-1051.

<sup>exciv</sup> Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. "Safe and Convenient Procedure for Solvent Purification" *Organometallics* **1996**, 15, 1518-1520.

<sup>excvi</sup> Still, W. C.; Kahn, M.; Mitra, A. "Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution" *J. Org. Chem.* **1978**, 43, 2923-2925.